

UNIVERSIDADE FEDERAL DO PARANÁ

JAIRO MARLON CORRÊA

MÉTODO WARIMAX-GARCH NEURAL PARA PREVISÃO DE SÉRIES
TEMPORAIS

CURITIBA
2015

JAIRO MARLON CORRÊA

MÉTODO WARIMAX-GARCH NEURAL PARA PREVISÃO DE SÉRIES
TEMPORAIS

Tese apresentada ao Curso de Pós-Graduação em Métodos Numéricos em Engenharia dos Setores de Tecnologia e de Ciências Exatas da Universidade Federal do Paraná, na Área de Concentração em Programação Matemática e na Linha de Pesquisa em Métodos Estatísticos Aplicados à Engenharia como requisito parcial à obtenção do título de Doutor em Métodos Numéricos em Engenharia.

Orientador: Prof. Dr. Anselmo Chaves Neto.

Coorientador: Prof. Dr. Luiz Albino Teixeira Júnior.

CURITIBA
2015

C824m

Corrêa, Jairo Marlon

Método Warimax-Garch Neural para previsão de séries temporais/ Jairo Marlon Corrêa. – Curitiba, 2015.

173 f. : il. color. ; 30 cm.

Tese - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-graduação em Métodos Numéricos em Engenharia, 2015.

Orientador: Anselmo Chaves Neto – Co-orientador: Luiz Albino Teixeira Júnior.

Bibliografia: p. 123-128.

1. Análise de séries temporais - Predição. 2. Barragens de concreto - Instrumentação. 3. Redes neurais (Computação). I. Universidade Federal do Paraná. II. Chaves Neto, Anselmo. III. Teixeira Júnior, Luiz Albino. IV. Título.

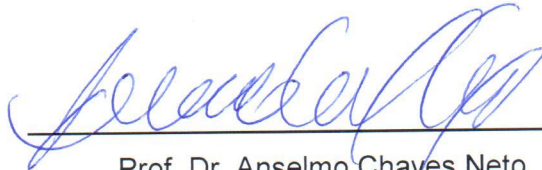
CDD: 519.55

TERMO DE APROVAÇÃO

JAIRO MARLON CORREA

MÉTODO WARIMAX-GARCH NEURAL PARA PREVISÃO DE SÉRIES TEMPORAIS

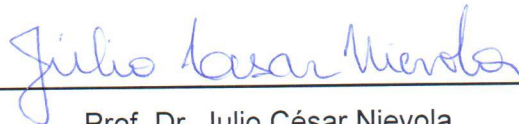
Tese aprovada como requisito parcial para obtenção do grau de doutor no Programa de Pós-Graduação em Métodos Numéricos em Engenharia, da Universidade Federal do Paraná, pela seguinte banca examinadora:



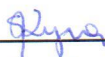
Prof. Dr. Anselmo Chaves Neto
Orientador – membro do PPGMNE/UFPR



Prof. Dr. Inácio Andruski Guimarães
Membro da UTFPR – Curitiba/PR



Prof. Dr. Julio César Nievola
Membro do PPGIA – PUC-PR



Prof.ª Dr.ª Sachiko Araki Lira
Membro do PPGMNE/UFPR

(participação por vídeo conferencia)
Prof. Dr. Reinaldo Castro Souza
Membro da PUC-RIO

Curitiba, 21 de dezembro de 2015.

Dedico este trabalho a Deus,
Responsável por toda a existência.

A minha esposa Keila,
que sempre permaneceu ao meu lado,
principalmente quando mais precisei.

Aos meus pais,
Jair e Margaret,
que sempre me apoiam e acreditam
nos meus estudos.

Aos meus irmãos, Greisy e Rodrigo.

AGRADECIMENTOS

A Deus pela proteção e bênçãos concedidas durante toda esta longa caminhada.

Aos professores Anselmo Chaves Neto e Luiz Albino Teixeira Junior pelas orientações, incentivos, persistências, ensinamentos e descontrações. Pela grande admiração que tenho a ambos, tanto profissional quanto pessoalmente.

A coordenadora professora Liliana Madalena Gramani, pelas cobranças motivadoras ao longo da caminhada. Por acreditar no potencial da nossa turma de Doutorado.

Aos meus amigos Samuel Bellido Rodrigues, Levi Lopes Teixeira, Tásia Hickmann e Lucas da Silva Ribeiro pelos momentos de estudos e descontração que compartilhamos.

Aos professores Arinei Carlos Lindbeck da Silva, Neida Maria Patias Volpi e Cassius Tadeu Scarpin pelos ensinamentos.

A minha esposa, por me incentivar nos estudos, por estar sempre ao meu lado e, principalmente, por entender a minha ausência quando se fez necessária.

Aos meus pais, Jair Corrêa e Margaret Carolina Albuquerque, aos meus irmãos Greisy Darling Corrêa e Jadson Rodrigo Corrêa pelo incentivo.

A Universidade Federal do Paraná, Universidade Estadual do Oeste do Paraná, Itaipu Binacional, CEASB e a todos que colaboraram com a viabilização do Doutorado.

Aos amigos e colegas do curso, que dividiram não só os momentos difíceis mas também os momentos de alegria e descontração ao longo do Doutorado.

A Universidade Tecnológica Federal do Paraná, campus Medianeira, a qual sou filiado. Aos colegas do Departamento de Matemática pelos momentos de descontração e produção.

SUMÁRIO

1	INTRODUÇÃO	17
1.1	PROBLEMA	23
1.2	OBJETIVOS	25
1.3	JUSTIFICATIVA	25
1.4	VALIDAÇÃO DO MÉTODO PROPOSTO	27
1.5	ESTRUTURA DA TESE	27
2	FUNDAMENTAÇÃO TEÓRICA	29
2.1	REVISÃO BIBLIOGRÁFICA	29
2.1.1	Análise <i>Wavelet</i>	29
2.1.2	Modelos ARIMA (<i>p, d, q</i>).....	33
2.1.3	Modelos ARIMAX-GARCH	36
2.1.4	Redes Neurais Artificiais	38
2.1.4.1	Componentes Básicas de uma RNA.....	41
2.1.4.1.1	Função de Ativação	41
2.1.4.1.2	Topologia.....	42
2.1.4.1.3	Treinamento Neural	43
2.1.4.1.4	Normalização dos Padrões.....	45
2.1.4.2	RNA's multicamadas <i>Feedforward</i> , com uma Camada Escondida	47
2.1.4.3	Geração de Previsões via Redes Neurais Artificiais.....	48
2.1.5	Combinação de Previsões	49
2.2	REVISÃO DA LITERATURA	51
3	MATERIAIS E MÉTODOS.....	56
3.1	MATERIAIS	56

3.2	MÉTODO PROPOSTO	61
4	RESULTADOS E DISCUSSÃO	68
4.1	EXPERIMENTO COMPUTACIONAL: MÉTODO WARIMAX-GARCH	68
4.1.1	Modelo ARIMA-GARCH	70
4.1.2	Modelo WARIMAX-GARCH: Série F05x	72
4.2	RESULTADOS COMPUTACIONAIS: MÉTODO WARIMAX-GARCH NEURAL.....	80
4.2.1	Método WARIMAX-GARCH NEURAL: Série F05y	80
4.2.1.1	Modelo ARIMA-GARCH	82
4.2.1.2	Modelo WARIMAX-GARCH NEURAL.....	83
4.2.2	Método WARIMAX-GARCH NEURAL: Série F13x	90
4.2.2.1	Modelo ARIMA-GARCH	93
4.2.2.2	Modelo WARIMAX-GARCH NEURAL.....	94
4.2.3	Método WARIMAX-GARCH NEURAL: Série F13y	100
4.2.3.1	Modelo ARIMA-GARCH	103
4.2.3.2	Modelo WARIMAX-GARCH NEURAL.....	104
4.2.4	Método WARIMAX-GARCH NEURAL: Série F19x	110
4.2.4.1	Modelo ARIMA-GARCH	113
4.2.4.2	Modelo WARIMAX-GARCH NEURAL.....	114
5	CONCLUSÃO E CONSIDERAÇÕES FINAIS	121
6	REFERÊNCIAS	123
7	APÊNDICES	129

LISTA DE SIGLAS

APE – *Absolute Percent Error*

ARCH – *Autoregressive Conditional Heteroskedasticity*

ARIMA – *Autoregressive Integrated Moving Average*

CL – Combinação Linear

DFA – Dikey-Fuller Aumentado

DW – Decomposição Wavelet

EQM - Erro Quadrático Médio

FAC – Função de Autocorrelação

FACP – Função de Autocorrelação Parcial

GARCH – *Generalized Autoregressive Conditional Heteroskedasticity*

GARCH-M – *GARCH in the Mean*

GED – Distribuição do Erro Generalizada

MAE – *Mean Absolute Error*

MAPE – *Mean Absolute Percent Error*

MLP – *Multi Layer Perceptron*

MLV - *Maximum Likelihood*

MQO – Mínimos Quadrados Ordinários

MSE – *Mean Squared Error*

MV – Máxima Verossimilhança

R^2 – Coeficiente de Determinação

RNA – Rede Neural Artificial

RMSE – *Root Mean Squared Error*

SARIMA – ARIMA Sazonal

VEW - Variável Exógena *Wavelet*

LISTA DE FIGURAS

Figura 1 – Tomada d'água de uma usina hidrelétrica	18
Figura 2 – Turbina de uma usina hidrelétrica em funcionamento.....	19
Figura 3 – Mapa de barragens da Usina Hidrelétrica de Itaipu	21
Figura 4 - Gráficos das funções escala e <i>Wavelet</i> de Haar	30
Figura 5 - Esquematização simplificada de um neurônio biológico	39
Figura 6 - Arquitetura básica de um neurônio artificial	40
Figura 7 - RNA MLP <i>feedforward</i>	48
Figura 8 - Ilustração da dinâmica de previsão, um passo à frente, de uma série temporal, com estrutura de autodependência igual a 4, através de uma RNA	49
Figura 9 – Instrumentação.....	57
Figura 10 - Estação total TM30	58
Figura 11 – Prismas de precisão localizados na barragem principal.....	59
Figura 12 – Rede Geodésica de Itaipu.....	60
Figura 13 - Fluxograma com etapas do método proposto	66
Figura 14 – Geração das variáveis exógenas híbridas	67
Figura 15 – Série temporal do Deslocamento do bloco F05x.....	68
Figura 16 – FAC e FACP dos resíduos ordinários padronizados	71
Figura 17 – FAC e FACP dos resíduos quadráticos padronizados	71
Figura 18 – Componentes <i>Wavelet</i> da amostra de treino da série F05x.....	72
Figura 19 - FAC e FACP dos resíduos ordinários padronizados.....	75
Figura 20 - FAC e FACP dos resíduos quadráticos padronizados	75
Figura 21 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH e ARIMA-GARCH	77
Figura 22 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH e RNA	78
Figura 23 – Evolução temporal do intervalo de confiança para ARIMA-GARCH	79
Figura 24 – Evolução Temporal do Intervalo de Confiança para WARIMAX-GARCH.....	79
Figura 25 – Série temporal do deslocamento do bloco F05y	80
Figura 26 – Série temporal do deslocamento do bloco F05y	81
Figura 27 – FAC e FACP dos resíduos ordinários padronizados	83

Figura 28 – FAC e FACP dos resíduos quadráticos padronizados	83
Figura 29 – Componentes <i>Wavelet</i> da amostra de treino da série F05y.....	84
Figura 30 – FAC e FACP dos resíduos ordinários padronizados	86
Figura 31 – FAC e FACP dos resíduos quadráticos padronizados	87
Figura 32 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH Neural e ARIMA-GARCH.....	88
Figura 33 – Evolução temporal do intervalo de confiança para ARIMA-GARCH	89
Figura 34 – Evolução Temporal do Intervalo de Confiança para WARIMAX- GARCH Neural	90
Figura 35 – Série temporal do deslocamento do bloco F13x	91
Figura 36 – Série temporal do deslocamento do bloco F13x	92
Figura 37 – FAC e FACP dos resíduos ordinários padronizados	94
Figura 38 – FAC e FACP dos resíduos quadráticos padronizados	94
Figura 39 – Componentes <i>Wavelet</i> da amostra de treino da série F13x.....	95
Figura 40 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH Neural e ARIMA-GARCH.....	98
Figura 41 – Evolução temporal do intervalo de confiança para ARIMA-GARCH	99
Figura 42 – Evolução Temporal do Intervalo de Confiança para WARIMAX- GARCH Neural	100
Figura 43 – Série temporal do deslocamento do bloco F13y	101
Figura 44 – Série temporal do deslocamento do bloco F13y	102
Figura 45 – FAC e FACP dos resíduos ordinários padronizados	104
Figura 46 – FAC e FACP dos resíduos quadráticos padronizados	104
Figura 47 – Componentes <i>Wavelet</i> da amostra de treino da série F13y.....	105
Figura 48 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH Neural e ARIMA-GARCH.....	108
Figura 49 – Evolução temporal do intervalo de confiança para ARIMA-GARCH	109
Figura 50 – Evolução Temporal do Intervalo de Confiança para WARIMAX- GARCH Neural	110
Figura 51 – Série temporal do deslocamento do bloco F19x	111
Figura 52 – Série temporal do Deslocamento do bloco F13x.....	112
Figura 53 – FAC e FACP dos resíduos ordinários padronizados	114
Figura 54 – FAC e FACP dos resíduos quadráticos padronizados	114

Figura 55 – Componentes <i>Wavelet</i> da amostra de treino da série F19x.....	115
Figura 56 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH Neural e ARIMA-GARCH.....	118
Figura 57 – Evolução temporal do intervalo de confiança para ARIMA-GARCH	119
Figura 58 – Evolução Temporal do Intervalo de Confiança para WARIMAX- GARCH Neural	120
Figura 59 – Evolução temporal do erro APE para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA.....	137
Figura 60 – Evolução temporal do erro APE para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA.....	146
Figura 61 – Evolução temporal do erro APE para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA.....	155
Figura 62 – Evolução temporal do erro APE para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA.....	163
Figura 63 – Evolução temporal do erro APE para WARIMAX-GARCH Neural e ARIMA-GARCH e RNA.....	172

LISTA DE TABELAS

Tabela 1 - Séries de deslocamentos dos blocos-chave da barragem principal da Usina Hidrelétrica de Itaipu com base nos pêndulos automatizados	60
Tabela 2 – Teste BDS dentro da amostra de teste para os resíduos ordinários do modelo ARIMA-GARCH.....	71
Tabela 3 – Valores dos erros MAPE e MAE, dentro e fora da amostra, para a série temporal F05X.....	76
Tabela 4 – Valores dos erros MAPE, MAE e RMSE, dentro e fora da amostra, para a série temporal F05Y	88
Tabela 5 – Valores dos erros MAPE, MAE e RMSE, dentro e fora da amostra, para a série temporal F13X	97
Tabela 6 – Valores dos erros MAPE, MAE e RMSE, dentro e fora da amostra, para a série temporal F13Y	107
Tabela 7 – Valores dos erros MAPE, MAE e RMSE, dentro e fora da amostra, para a série temporal F19X	117
Tabela 8 – Configurações das RNA-MLPs para as componentes wavelet.	141
Tabela 9 – Configurações das RNA-MLPs para as componentes <i>wavelet</i>	148
Tabela 10 – Configurações das RNA-MLPs para as componentes <i>wavelet</i>	157
Tabela 11 – Configurações das RNA-MLPs para as componentes <i>wavelet</i>	165

APÊNDICES

Apêndice I A: Teste Dikey-Fuller Aumentado (série temporal F05x).....	129
Apêndice I B: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (série temporal F05x).....	129
Apêndice I C: Modelo WARIMAX-GARCH e seus resultados estatísticos para validação (Série temporal F05x).	130
Apêndice I D: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX- GARCH Neural (F05x).....	136
Apêndice II A: Teste Dikey-Fuller Aumentado (série temporal F05y).....	139
Apêndice II B: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (série temporal F05y).....	139
Apêndice II C: Configurações da RNA (série temporal F05y).	140
Apêndice II D: Modelo WARIMAX-GARCH e seus resultados estatísticos para validação (F05y).	141
Apêndice II E: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX- GARCH Neural (F05y).....	145
Apêndice III A: Teste Dikey-Fuller Aumentado (F13x).	148
Apêndice III B: Configurações da RNA (série temporal F13x).....	148
Apêndice III C: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (F13x).	149
Apêndice III D: Modelo WARIMAX-GARCH Neural e seus resultados estatísticos para validação (F13x).	150
Apêndice III E: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX-GARCH Neural (F13x).....	153
Apêndice IV A: Teste Dikey-Fuller Aumentado (F13y).....	157
Apêndice IV B: Configurações da RNA (série temporal F13y).	157
Apêndice IV C: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (série temporal F13y).....	158

Apêndice IV D: Modelo WARIMAX-GARCH Neural e seus resultados estatísticos para validação (série temporal F13y).....	159
Apêndice IV E: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX-GARCH Neural (F13y).....	162
Apêndice V A: Teste Dikey-Fuller Aumentado (série temporal F19x).	165
Apêndice V B: Configurações da RNA (série temporal F19x).	165
Apêndice V C: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (série temporal F19x).....	166
Apêndice V D: Modelo WARIMAX-GARCH Neural e seus resultados estatísticos para validação (série temporal F19x).....	167
Apêndice V E: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX-GARCH Neural (F13y).	171

RESUMO

A proposta deste trabalho é apresentar uma nova metodologia híbrida WARIMAX-GARCH Neural para a previsão pontual e intervalar de séries temporais estocásticas. Fundamentalmente, é aplicada a decomposição *Wavelet* em séries históricas compostas por registros de monitoramento de barragens e suas componentes de aproximação e detalhe, as quais são modeladas, individualmente, via ARIMA-GARCH e Redes Neurais Artificiais (RNA). A partir de então, são realizadas as previsões pontuais fora da amostra pelas técnicas de modelagem e os resultados são combinados linearmente. As componentes de aproximação e detalhe são completadas com as previsões combinadas e passam a ser utilizadas como variáveis de entrada (exógenas híbridas) na modelagem da série em estudo. Em cada série temporal é aplicada a metodologia WARIMAX-GARCH Neural e são realizadas as previsões pontuais e intervalares, sob a suposição de inovações gaussianas. As séries temporais utilizadas neste trabalho de tese foram as séries temporais dos deslocamentos horizontais de blocos da barragem principal da Usina Hidrelétrica de Itaipu, aferidas pelos pêndulos diretos automatizados. Os desempenhos preditivos alcançados pela metodologia proposta, em relação aos resultados obtidos pelas modelagens tradicionais ARIMA-GARCH e RNA, foram consideravelmente vantajosos. Nas comparações dos resultados obtidos através do modelo WARIMAX-GARCH Neural com métodos tradicionais, a redução do erro preditivo chegou a 91%.

Palavras-chave: Monitoramento de Barragens, Previsão de Séries Temporais, Modelos ARIMA, Modelos GARCH, Redes Neurais Artificiais, Decomposição *Wavelet*.

ABSTRACT

This research proposes a new WARIMAX-GARCH Neural hybrid methodology for point and interval prediction of stochastic time series. Fundamentally, it is applied the wavelet decomposition on the time series made of monitoring data and its approximation and detail components were modeled by ARIMA-GARCH and Artificial Neural Networks (ANN). Thereafter, the point forecasts are performed out the sample by both modeling techniques and these results are combined linearly. The approximation and detail components are completed with the combined forecasts and are used as input variables (hybrid exogenous) in the modeling time series under study. In each time series is applied the WARIMAX-GARCH Neural methodology and are made the point and interval forecast, under the assumption of Gaussian innovations. The time series used in this research were the time series of horizontal displacements of the main dam blocks of Itaipu hydroelectric plant, measured by automated direct pendulums. The predictive performances achieved by the proposed method compared to the results obtained by traditional modeling ARIMA-GARCH and RNA were considerably advantageous. Comparing the results obtained by WARIMAX-GARCH Neural model to traditional methods there was a reduction of up to 91% of the predictive error.

Keywords: Dams Monitoring, Forecast Time Series, ARIMA Models, GARCH Models, Artificial Neural Networks, Wavelet Decomposition.

1 INTRODUÇÃO

O Brasil é um dos países de maior potencial energético do mundo, sendo que a hidráulica se destaca entre as fontes energéticas disponíveis a hidráulica. Segundo a publicação Balanço Energético Nacional (BEN), edição 2015, ano base 2014, o Brasil dispõe de uma matriz elétrica de origem renovável, com destaque para a geração hidráulica que corresponde a 65,2% da oferta interna. Fato é que a energia hidráulica é a principal fonte geradora de energia elétrica no país, e continuará sendo por muitos anos. Somente a bacia do rio Paraná é responsável por quase 50% da capacidade hidrelétrica instalada no Brasil. Das fontes renováveis é considerada a de maior representatividade na matriz energética mundial por possuir tecnologias aprimoradas, basta considerar as usinas hidrelétricas já construídas (ANEEL, 2008).

A importância do uso da energia hidráulica predomina desde os primórdios das civilizações devido à disponibilidade de recursos, facilidade de aproveitamento e, sobretudo, seu caráter renovável. Além disso, ela contribuiu para o desenvolvimento econômico de forma expressiva como, por exemplo, em atividades agrícolas, de serviços, comerciais e, também, na melhoria do conforto das residências e condição de vida das pessoas. Atualmente representa cerca de 20% de toda a eletricidade gerada no mundo e é aproveitada como principal fonte na produção de energia elétrica em mais de 30 países, dentre os quais se destacam como maiores produtores mundiais: Canadá, China, Brasil, Estados Unidos e Rússia.

A energia elétrica produzida em uma hidrelétrica, do tipo gravidade aliviada, é obtida a partir da energia cinética. Para obter este tipo de energia são construídas as barragens que são responsáveis pelo represamento do rio e, com isso, há o acúmulo de água (montante) para fornecimento do desnível e, conseqüentemente, a sua queda. A barragem interrompe a água, o que força a sua passagem pela captação, denominada tomada d'água (Figura 1), que é

levada por condutos forçados para as turbinas. A energia potencial é responsável pelo acionamento ou giro das turbinas (Figura 2) que são postas em rotação a partir do recebimento da massa de água em suas pás. O movimento rotatório das turbinas é convertido em energia elétrica a partir de um gerador. A água não utilizada na produção de energia elétrica é desviada para o leito do rio por meio do vertedouro, que consiste de comportas responsáveis pelo controle do nível máximo do lago.

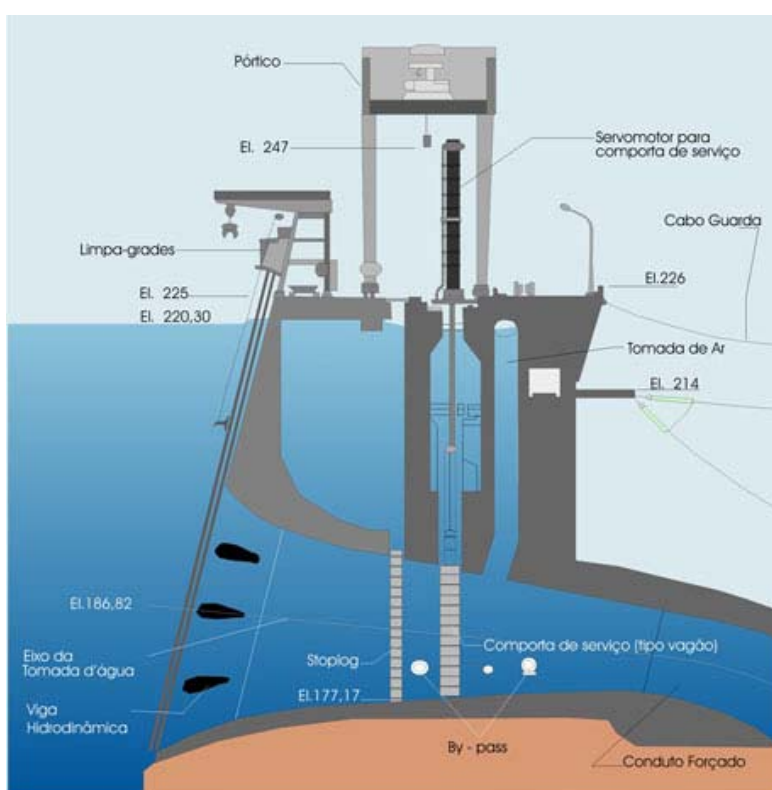


Figura 1 – Tomada d'água de uma usina hidrelétrica



Figura 2 – Turbina de uma usina hidrelétrica em funcionamento

A Usina Hidrelétrica de Itaipu é o resultado do comprometimento de dois países vizinhos, Brasil e Paraguai, cuja origem das negociações ocorreu em 1966 e, em decorrência, foi assinado o “Tratado de Itaipu”. Este tratado é composto por 25 artigos que regulamentam o aproveitamento de recursos hídricos do rio Paraná, a participação equilibrada na administração, a execução do projeto e o fornecimento equivalente de mão de obra.

A empresa criada pelo tratado é a Itaipu Binacional. Trata-se de um dos maiores projetos hidrelétricos do mundo, considerando-se a produção anual. Está localizado no rio Paraná a 14 km à montante da ponte internacional que liga a cidade de Foz do Iguaçu, no Brasil, à Ciudad del Este no Paraguai. O projeto da barragem da Itaipu foi bastante criterioso tanto em termos de investigação da rocha de fundação quanto dos materiais utilizados na construção das barragens, seguindo um minucioso controle de qualidade e tendo a instalação de milhares de instrumentos cujas leituras são feitas desde a época da construção. As campanhas de leituras e medições têm uma periodicidade que varia de semanal a semestral, conforme o tipo de instrumento. Além disso, cerca de 10% dos principais instrumentos foram automatizados para se obter leituras mais frequentes, permitindo sistemas de

alerta. A análise dos dados é feita por especialistas em segurança de barragens que emitem pareceres técnicos semestralmente e a cada quatro anos todos os resultados e estudos são apresentados para uma junta de consultores internacionais que avalia todos os aspectos relacionados à segurança de barragens.

Diante disso, as barragens são obras que requerem atenção às diversas condições de segurança estrutural e operacional, a fim de evitar algum dano em suas estruturas ou até mesmo ao meio ambiente ou perdas de vidas humanas. A identificação de problemas e a execução de reparos diminuem o fator risco. Porém, esses estudos requerem muita atenção para a determinação de soluções adequadas. Tendo isso como base, a Usina de Itaipu foi concebida com um plano de auscultação que conta com mais de 2400 instrumentos e um sistema de drenagem com mais de 5000 drenos, instalados em suas estruturas e fundações, formando um banco de dados de mais de 30 anos. Alguns blocos são designados blocos-chave, pois são dotados de maior quantidade de instrumentos devido a sua importância para a região monitorada e características prédefinidas em projeto.

Há diversos tipos de barragens que compõem, estruturalmente, a Usina Hidrelétrica de Itaipu, conforme mapa da Figura 3. A barragem principal é de concreto, do tipo gravidade aliviada e é composta por 18 blocos, onde estão as tomadas d'água de 16 unidades geradoras. Cada turbina possui potência nominal de 715MW que equivale a uma potência nominal do gerador de 700MW. No total, são 20 unidades geradoras, das quais 4 estão inseridas na estrutura de desvio. Ao giro de 85,7 rpm (rotações por minuto) e operando na frequência 60HZ as turbinas produzem energia para o sistema brasileiro, enquanto ao giro de 71,4 rpm e na frequência de 50HZ, é produzida a energia elétrica destinada ao sistema paraguaio (ITAIPU, 2009).

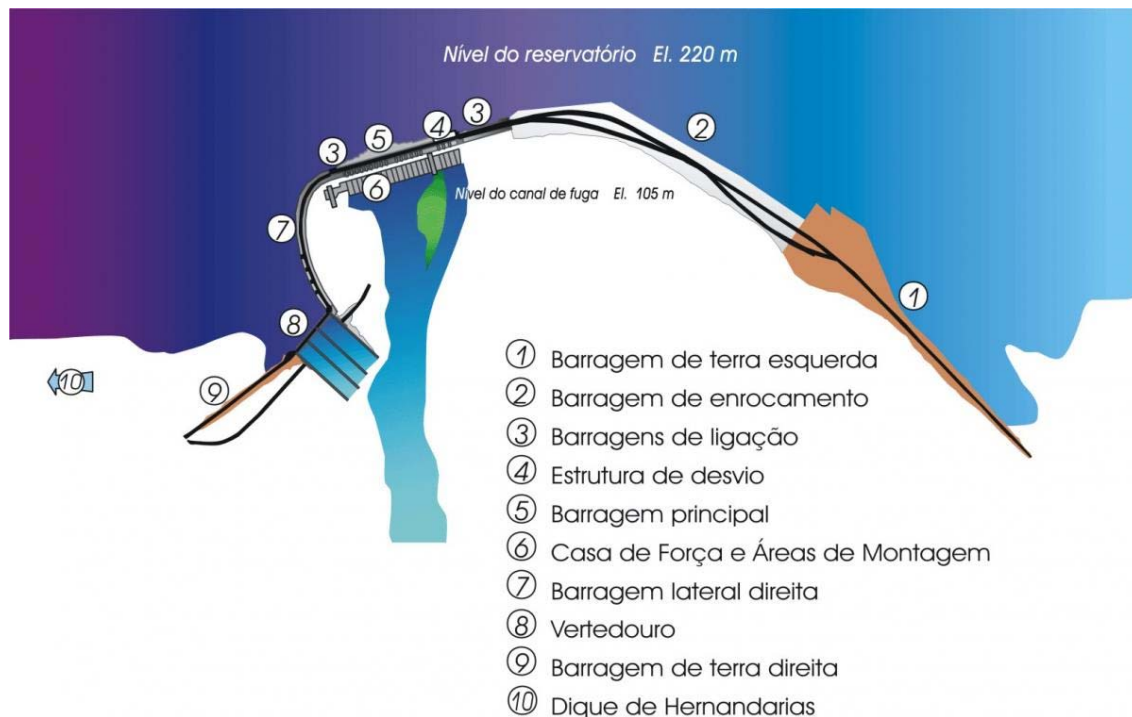


Figura 3 – Mapa de barragens da Usina Hidrelétrica de Itaipu
Fonte: Itaipu, 2015c

Em 2005, como parte da modernização da usina, foi implantado um sistema de aquisição de dados automáticos (ADAS), onde 10% dos instrumentos mais importantes foram automatizados, passando a ter leituras a cada 30 minutos, além da leitura manual já existente.

A água do reservatório exerce uma pressão na face montante da barragem e, também, exerce uma subpressão ou empuxo hidrostático nas partes inferiores dos blocos, que cria um efeito contrário ao que a sua própria massa exerce sobre a fundação. Devido a essas duas forças, a barragem tem uma tendência a apresentar pequenos deslocamentos horizontais, principalmente no sentido montante-jusante. Em contrapartida, as forças estabilizantes ou a favor da segurança são o peso próprio da estrutura e o atrito entre a barragem e a rocha de fundação. De acordo com Itaipu (2009), foram instalados dois tipos de pêndulos (ver Figura 9, página 57) em trechos da Usina Hidrelétrica de Itaipu: o pêndulo direto e o invertido. O pêndulo direto é responsável em monitorar os deslocamentos horizontais de pontos dos blocos instrumentados da barragem em determinadas cotas, em relação à fundação

da estrutura. Já a função do pêndulo invertido é monitorar os deslocamentos da fundação da barragem em relação ao contato concreto/rocha ou a alguma cota determinada no bloco de concreto.

É fato que a barragem acumulou um pequeno deslocamento ao longo de 32 anos desde o enchimento do reservatório. A análise dos dados dos pêndulos mostra que o deslocamento descreve um movimento que gera uma curva cujo comportamento é cíclico, influenciado pela temperatura ambiente. Alguns blocos acumularam um deslocamento com aproximadamente 13mm em relação à fundação, considerando o período de tempo desde o enchimento do lago até os dias atuais (ano 2015). Vale salientar que esses deslocamentos são considerados pequenos e estão dentro dos limites de segurança estabelecidos no projeto. Mesmo assim, a equipe técnica da Itaipu vem acompanhando periodicamente as leituras de todos os instrumentos e incentivando estudos para aplicação de ferramentas matemáticas para atuar preventivamente ante quaisquer riscos potenciais.

As aferições geradas pelos pêndulos formam um banco de dados que contém agrupamentos numérico-temporais. De acordo com Morettin e Toloí (2006), uma série temporal é qualquer conjunto de observações coletadas sequencialmente no decorrer do tempo. Basicamente, uma série temporal é composta por quatro elementos: tendência, que diz respeito ao sentido do deslocamento da série durante um longo espaço de tempo; ciclo, que são os movimentos ondulatórios que a série temporal apresenta a cada intervalo de tempo; sazonalidade, que são os efeitos que a série temporal apresenta em decorrência das variações climáticas e o ruído aleatório, que são os dados que não podem ser modelados, pois são gerados por influências externas ao processo, é também chamado de erro.

As informações pertinentes ao comportamento estrutural de uma barragem em uma usina hidrelétrica, por exemplo, são avaliadas por meio das leituras e análise dos dados de instrumentação. Estatisticamente, a análise gráfica dos dados temporais de um instrumento na busca da compreensão da sazonalidade, ciclos, bem como o uso de estatísticas como média, variância,

amplitude, valores máximos e mínimos históricos, podem não ser suficientes para se criar planos de ações ou criar referências para tomada de decisão. Nesta tese de Doutorado, propõe-se a implementação de uma metodologia estatística alternativa, híbrida, de previsão de séries temporais, capaz de mapear informações presentes na série temporal, como volatilidade, estruturas de autodependência linear e não linear, com aplicações em séries temporais reais. Estas características estão presentes nas séries de deslocamentos horizontais de blocos da barragem principal de Itaipu, aferidos por meio dos pêndulos automatizados que monitoram o comportamento estrutural da barragem, como será visto adiante.

Desse modo, as projeções das séries temporais dos deslocamentos horizontais podem operar em benefício do desempenho estrutural da barragem da Usina Hidrelétrica de Itaipu, uma vez que estas informações permitem aos engenheiros e à equipe técnica adotar medidas corretivas em planos de execução ou tomada de decisões já existentes.

1.1 Problema

A segurança de barragens é um assunto de grande relevância e complexidade. Inspeções regulares de uma barragem são uma das formas de monitoramento da sua saúde estrutural. Operações adequadas, manutenções e inspeções periódicas, planos de emergência frente a situações de risco e instrumentação suficiente são algumas estratégias adotadas que aumentam a sua segurança e vida útil.

O deslocamento estrutural de uma barragem é uma das informações de maior relevância para técnicos e engenheiros, uma vez que pequenos deslizamentos podem gerar minúsculas fissuras, resultando no enfraquecimento estrutural da barragem. Um dos instrumentos mais robusto, preciso e de baixo custo, responsável por mensurar o deslocamento da crista da barragem, é o pêndulo (SILVEIRA, 2013).

Diante deste fato, as séries temporais dos deslocamentos horizontais de blocos-chave da barragem principal da Usina Hidrelétrica de Itaipu, local onde

estão instaladas as unidades geradoras de energia, são aferidas pelos pêndulos diretos e invertidos. As previsões das séries temporais destes deslocamentos permitem a projeção, ao longo do tempo (dias e até meses), do comportamento estrutural da barragem.

As séries dos deslocamentos dos blocos da barragem principal de Itaipu, que foram utilizadas neste estudo, apresentam estruturas de dependência serial linear e não linear, bem como volatilidade (variância do erro não constante). Para a modelagem destas séries, foi concebida uma metodologia que apresenta vantagens preditivas comparada aos métodos tradicionais de previsão de séries de tempo, como os modelos ARIMA-GARCH e as Redes Neurais Artificiais.

Para modelagem das séries temporais que apresentam tais estruturas, propõe-se uma metodologia eficiente que aborde estas informações temporais. As etapas da metodologia proposta são:

- (i) Decomposição da série temporal original, usando a transformada *Wavelet*;
- (ii) Modelagem linear e da volatilidade (ARIMA-GARCH) e não linear (RNA) de cada uma das componentes *Wavelet*;
- (iii) Geração das previsões de cada componente *Wavelet* pelo método proposto em (ii);
- (iv) Combinação linear das previsões obtidas em (iii) e uso dessas variáveis como variáveis exógenas sintéticas ou artificiais;
- (v) Modelagem da série temporal original utilizando as séries obtidas em (iv) como componentes exógenas;
- (vi) Geração de previsões pontuais e intervalares da série subjacente.

As previsões geradas por meio da metodologia proposta trarão informações do deslocamento horizontal e revelarão as características e tendências do movimento que melhor se aproxima da realidade. Com isso, a equipe técnica e os engenheiros poderão adotar melhores estratégias em suas medidas preventivas e corretivas nos planos de execução que estes resultados possibilitarem. O monitoramento eficiente do deslocamento horizontal com as

previsões a vários passos à frente poderá prolongar a saúde estrutural e aumentar a vida útil da barragem.

1.2 Objetivos

Objetivo Geral

Analisar séries temporais de deslocamento de blocos da barragem principal, da Usina Hidrelétrica de Itaipu, e propor um método alternativo de previsão ARIMA-GARCH, a partir da decomposição *Wavelet*, com a geração de intervalos de confiança para as previsões pontuais obtidas pelo método.

Objetivos Específicos

- Propor um novo método de previsão chamado WARIMAX-GARCH Neural;
- Modelar séries de deslocamento horizontal de blocos-chave da barragem da Usina Hidrelétrica de Itaipu;
- Demonstrar empiricamente a eficácia do uso da decomposição *Wavelet* na previsão de séries temporais no presente estudo de caso;
- Analisar os ganhos preditivos obtidos com a combinação de previsores lineares e não lineares;
- Comparar os resultados alcançados com o método proposto aos obtidos por métodos de previsão tradicionais ARIMA-GARCH e RNA;
- Obter previsões pontuais e intervalares das séries temporais modeladas por meio do modelo desenvolvido, sob a suposição de resíduos gaussianos.

1.3 Justificativa

A decomposição de uma série temporal em componentes ortogonais (uma componente de aproximação e r componentes de detalhe), convencionada pela teoria *Wavelet*, possibilita a compreensão dos sinais nas

suas partes constituintes, ou seja, fornece uma melhor descrição do sinal de uma forma grosseira (componente de aproximação) e de uma forma detalhada desde os menos delicados aos mais finos (componentes de detalhes) (Lima, 2002). Com o uso desta técnica, em diversos trabalhos como Seo *et al.* (2015), Karthikeyan e Nagesh Kumar (2013), An *et al.* (2011), Benaouda *et al.* (2006), Jammazi e Aloui (2012), Soltani e Soltani (2002), Reboredo e Rivera-Castro (2013), Ramsey *et al.* (2010), esteve presente o uso da decomposição *Wavelet*. Essa inserção forneceu melhor representação do sinal em seus níveis de detalhes presentes. Teixeira Júnior *et al.* (2012) apresentaram um método que se compõe da combinação de uma rede neural artificial (RNA) e a decomposição *Wavelet* para gerar previsões de curto prazo da série temporal horária da radiação solar média (W/m^2). Nesse trabalho, os autores obtiveram resultados melhores que os obtidos apenas por RNA isoladamente.

A técnica de combinação linear ou não linear de previsões procedentes de métodos distintos leva a ganhos preditivos significativos, de maneira geral. Newbold e Granger (1974), Winkler e Makridakis (1983) e Teixeira Junior (2013) obtiveram a redução do erro preditivo com o uso da técnica de combinação de previsões. Segundo Werner e Ribeiro (2006) e Wallis (2011) não há uma melhor regra de combinação de métodos de previsão. Empiricamente, os trabalhos na área mostram que a combinação de previsões associada à decomposição *Wavelet* tem mostrado um ganho na análise e previsões de séries temporais. Trata-se, portanto, de uma área de pesquisas abertas a discussões e estabelecimentos de novos métodos e, nesse caso, um dos assuntos discutidos nesse trabalho.

Diante desses desafios, o presente estudo propõe um método para diagnosticar, também, o comportamento da barragem em virtude do deslocamento horizontal existente. Em consequência de um pequeno deslocamento horizontal há o comprometimento da estrutura toda, resultando no enfraquecimento da sustentação e ocorrência de fissuras. O controle dos deslocamentos horizontais de uma barragem permite um bom monitoramento e intervenções em eventuais anomalias da barragem ao longo de sua vida.

Dado que alguns aspectos ainda são desconhecidos, devido ao grande porte da Usina e a imensa quantidade de leituras efetuadas pelos instrumentos, os resultados obtidos com a aplicação de métodos irão mensurar o deslocamento horizontal dos blocos da barragem. Os resultados obtidos permitirão aos meios técnicos da Itaipu, por meio de um processo mais bem estruturado, a escolha de estratégias mais apropriadas para alcançar determinados objetivos no sentido de obter os melhores resultados.

1.4 Validação do Método Proposto

A avaliação do desempenho do método proposto será efetivada por meio da modelagem de séries temporais com o uso de métodos convencionais individuais como ARIMA-GARCH e RNA. Os resultados destes métodos convencionais serão comparados com os resultados do método proposto, onde serão avaliados os erros de previsão com as estatísticas RMSE (*Root-Mean-Square Error*), MAPE (*Mean Absolute Percent Error*) e MAE (*Mean Absolute Error*).

1.5 Estrutura da Tese

O presente trabalho está organizado em seis capítulos: (1) Introdução; (2) Fundamentação Teórica; (3) Materiais e Métodos; (4) Resultados e Discussão; (5) Conclusão e Considerações Finais. No capítulo 2 são explanados os conceitos e descrições de técnicas a serem usadas na pesquisa, tais como: Modelos de Box e Jenkins, modelos ARIMAX-GARCH; Decomposição *Wavelet*, RNA-MLP (Redes Neurais Artificiais *Multilayer Perceptron*). No capítulo 3 são apresentados o material (séries temporais disponibilizadas pela Itaipu Binacional) e a descrição do método proposto, designado por WARIMAX-GARCH Neural. Já no capítulo 4 são apresentados e discutidos os resultados obtidos com o emprego de algumas das técnicas presentes no método descrito no capítulo 3. No capítulo 5 estão as Considerações Finais e a Conclusão.

Finalmente, há o capítulo 6 com as Referências e o capítulo 7 dedicado aos Apêndices.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são tratados os aspectos básicos atinentes às seguintes abordagens: Análise *Wavelet* (seção 2.1.1); Modelos ARIMA (p, d, q) (seção 2.1.2); Modelos ARIMAX-GARCH (seção 2.1.3); Redes Neurais Artificiais (seção 2.1.4); Combinação de Previsões (seção 2.1.5).

2.1 Revisão bibliográfica

2.1.1 Análise *Wavelet*

Os primeiros estudos feitos sobre *Wavelet* ocorreram em 1909, por Haar (1910). O estudo da teoria *Wavelet* ficou no anonimato por um período longo, sendo que em 1985, Stephanie Mallat deu um grande impulso a este tema por meio de seu trabalho na área de processamento digital de imagens. O estudo seguiu com Mallat (1989) e com Daubechies (1990), que criou um conjunto de bases ortonormais de *Wavelet* suaves, com suportes compactos. Estes últimos trabalhos são os alicerces das aplicações atuais de *Wavelet*.

Intuitivamente, pode-se dizer que a função *Wavelet* permite a decomposição ou representação de uma série temporal, originalmente descrita num domínio de tempo, para que esta possa ser analisada em diferentes escalas de frequência de tempo. Em outras palavras, as *Wavelets* compõem uma técnica matemática que permite a decomposição ortogonal de funções de energia finita (isto é, quadraticamente somáveis ou integráveis) (LIMA, 2002).

Segundo Graps (1995), a análise de Fourier utiliza funções bases (senos e cossenos) para decompor o sinal em ondas senoidais de várias bandas de frequências, enquanto na análise *Wavelet* um sinal $f(x)$ é decomposto a partir de informações de escala (compressão e dilatação) e translação de uma função *Wavelet*. Vale salientar que as *Wavelet* permitem a análise da função em relação ao tempo, o que não é possível na análise Fourier.

O mapeamento das componentes de alta frequência de um sinal $f(x)$ é descrito pela função *Wavelet* $\omega(x)$ e as de baixa frequência obtidas pela função

escala $\phi(x)$. Entre as famílias de funções *Wavelet* vigentes na literatura, a mais simples, e também a mais antiga, é a de Haar, introduzida por Alfred Haar (HAAR, 1910), a qual é definida em (1)

$$\omega(x) = \begin{cases} 1, & \text{se } x \in [0, 1/2) \\ -1, & \text{se } x \in [1/2, 1) \\ 0, & \text{caso contrário} \end{cases} \quad (1)$$

A definição da função escala da família de Haar é dada em (2)

$$\phi(x) = \begin{cases} 1, & \text{se } x \in [0, 1) \\ 0, & \text{caso contrário} \end{cases} \quad (2)$$

Na Figura 4 ((a) e (b)) estão representados os gráficos das funções escala e *Wavelet* de Haar, respectivamente.

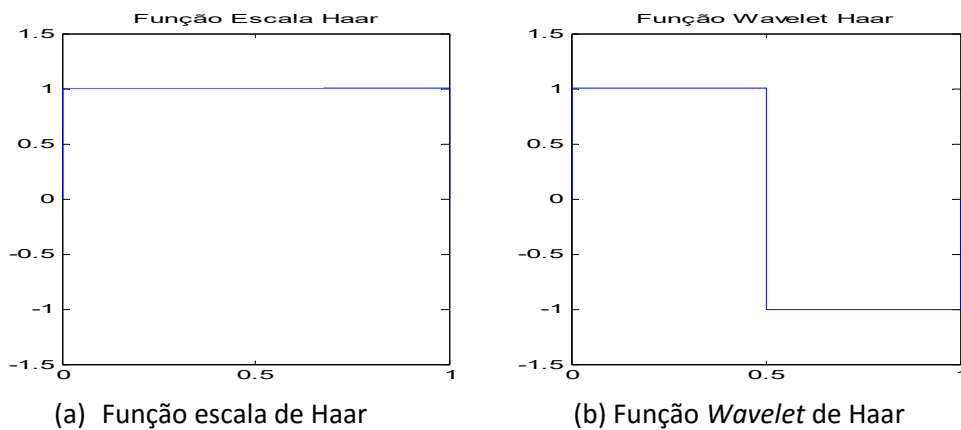


Figura 4 - Gráficos das funções escala e *Wavelet* de Haar

Matematicamente, as funções *Wavelet* são elementos do espaço de Hilbert¹ l^2 (Kubrusly, 2001) constituído por funções quadraticamente somáveis. A função $\omega(\cdot)$ pertencente a l^2 é chamada de função *Wavelet* se, e somente se, as funções $\omega_{m,n}(\cdot) = 2^{\frac{m}{2}}\omega(2^m(\cdot) - n)$, onde $m, n \in \mathbb{Z}$, formarem uma base ortonormal para o espaço de Hilbert l^2 ; segundo Levan e Kubrusly (2003), o parâmetro m é chamado de parâmetro de escala e n de parâmetro de translação. Desta forma, a função $f(\cdot) \in l^2$ admite sua expansão ortogonal por

¹ Espaços produto interno completos

meio de uma série de Fourier, em termos de uma base ortonormal *Wavelet* $\{\omega_{m,n}(\cdot)\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$ para l^2 , definida como em (3).

$$f(\cdot) := \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f(\cdot), \omega_{m,n}(\cdot) \rangle \omega_{m,n}(\cdot) \quad (3)$$

De acordo com Kubrusly e Levan (2006), o subespaço fechado (4)

$$W_m(\omega) := \left(\text{span} \{ \omega_{m,n}(\cdot) \}_{n \in \mathbb{Z}} \right)^- \quad (4)$$

contido em l^2 é chamado de subespaço de detalhes, na escala m . Segundo Levan e Kubrusly (2003), a projeção ortogonal de $f(\cdot)$, sobre o subespaço de detalhes $W_m(\omega)$, é definida pela soma parcial definida em (5).

$$f_{W_m(\omega)}(\cdot) := \sum_{n \in \mathbb{Z}} \langle f(\cdot), \omega_{m,n}(\cdot) \rangle \omega_{m,n}(\cdot) \quad (5)$$

Segundo Daubechies (1988), a projeção $f_{W_m(\omega)}(\cdot)$ pode ser referida como uma componente de detalhe *Wavelet* de $f(\cdot)$, na escala m , sobre $W_m(\omega)$.

A função $\phi(\cdot) \in l^2$ é chamada de função (*Wavelet*) *escala* se, e somente se, a coleção de funções $\phi_{m,n}(\cdot) = 2^{\frac{m}{2}} \phi(2^m(\cdot) - n)$, onde $m, n \in \mathbb{Z}$, são tais que $\langle \phi_{l,i}(\cdot), \phi_{j,k}(\cdot) \rangle = 0$ sempre que $l = j$ e $i \neq k$, e $\langle \phi_{l,i}(\cdot), \phi_{j,k}(\cdot) \rangle \neq 0$, se caso contrário.

De acordo com Mallat (1989), o subespaço fechado $V_m(\phi)$ de l^2 , onde $V_m(\phi) := \left(\text{span} \{ \phi_{m,n}(\cdot) \}_{n \in \mathbb{Z}} \right)^-$ é chamado de subespaço de aproximação, na escala m . Segundo Levan e Kubrusly (2003), a projeção ortogonal de $f(\cdot)$ sobre o subespaço de aproximação $V_m(\phi)$ é definida pela soma parcial, definida em (6).

$$f_{V_m(\phi)}(\cdot) := \sum_{n \in \mathbb{Z}} \langle f(\cdot), \phi_{m,n}(\cdot) \rangle \phi_{m,n}(\cdot) \quad (6)$$

De acordo com Daubechies (1988), a projeção $f_{V_m(\phi)}(\cdot)$ pode ser referida como uma componente de aproximação *Wavelet* de $f(\cdot)$, na escala m , sobre $(V_m(\phi), \langle ; \rangle)$.

Uma transformada *Wavelet* sobre l^2 é definida por um produto interno $\langle \cdot, \cdot \rangle$ entre uma função $f(\cdot) \in l^2$ e uma função *Wavelet* $\omega_{m,n}(\cdot)$ ou uma função escala $\phi_{m,n}(\cdot)$, onde $(m,n) \in \mathbb{Z} \times \mathbb{Z}$. De acordo com Morettin e Tolo (2006), as transformadas *Wavelet* podem ser agrupadas em dois conjuntos disjuntos: o de coeficientes de detalhe, denotado por $\{d_{m,n}\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$, e o de coeficientes de aproximação, denotado por $\{a_{m,n}\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$, onde um coeficiente de detalhe $d_{m,n}$ e um de aproximação $a_{m,n}$ são definidos, respectivamente, em (7) e (8).

$$d_{m,n} := \langle f(\cdot), \omega_{m,n}(\cdot) \rangle = \sum_{t \in \mathbb{Z}} f(t) \omega_{m,n}(t) \quad (7)$$

$$a_{m,n} := \langle f(\cdot), \phi_{m,n}(\cdot) \rangle = \sum_{t \in \mathbb{Z}} f(t) \phi_{m,n}(t) \quad (8)$$

Baseado em Teixeira Júnior *et al.* (2015) e Kubrusly e Levan (2006), a expansão de um mapa $f(\cdot)$ sobre o espaço de Hilbert l^2 , em termos de uma base ortonormal *Wavelet*, é definida, genericamente, por

$$f(\cdot) = \sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot) + \sum_{m=m_0}^{+\infty} \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(\cdot) \quad (9)$$

onde: $\sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot)$ é a componente *Wavelet* de aproximação de nível m_0 do mapa $f(\cdot)$; $\sum_{m=m_0}^{+\infty} \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(\cdot)$ é a componente de detalhe de nível m do mapa $f(\cdot)$.

Tautologicamente, qualquer série temporal finita y_t ($t = 1, \dots, T$) pode ser interpretada como uma sequência infinita y_t ($t \in \mathbb{Z}$) em l^2 definido como $y_t = y_t$, se $t \in \{1, 2, \dots, T\}$; e $y_t = 0$, se $t \in \mathbb{Z} - \{1, 2, \dots, T\}$. Portanto, qualquer série temporal y_t ($t = 1, \dots, T$) finita, pode ser decomposta ortogonalmente pela decomposição *Wavelet* apresentada em (9).

Em termos práticos, uma vez que é possível modelar separadamente todas as componentes *Wavelets* geradas pela expansão (9), uma adaptação é necessária para obter um número finito de componentes. Então, de acordo com Donoho e Johnstone (1994) Donoho; Johnstone (1994), uma boa alternativa pode ser apresentada pela decomposição *Wavelet* de nível r da série de tempo y_t ($t = 1, \dots, T$), onde $r > 0$ e $r \in \mathbb{Z}$, é dada por

$$y_t = \tilde{y}_{A_{m_0},t} + \sum_{m=m_0}^{m_0+(r-1)} \tilde{y}_{D_m,t} + \varepsilon_t, \quad (10)$$

onde m_0 é o parâmetro de escala (que muitas vezes é assumido como sendo igual a r); ε_t é o vetor de erro da aproximação de y_t (que é geralmente assumido como sendo um vetor nulo); $\tilde{y}_{A_{m_0},t} := \sum_{n=0}^{2^{(M-m_0)}-1} a_{m_0,n} \phi_{m_0,n}(t)$ e $\tilde{y}_{D_m,t} := \sum_{n=0}^{2^{(M-m_0)}-1} d_{m,n} \omega_{m,n}(t)$, que são, respectivamente, as componentes de aproximação de nível m_0 e de detalhes de nível m e consistem, respectivamente de estimativas para $y_{A_{m_0},t}$ e $y_{D_m,t}$ em (9); M é um parâmetro que tem um valor inteiro tal que $T = 2^M$. Se T não é uma potência inteira de 2, muitas vezes se completa y_t ($t = 1, \dots, T$) com zeros de tal modo que o seu comprimento T é aumentado até a próxima potência inteira de 2. Este procedimento pode ser realizado porque os zeros que se adicionaram não afetam o cálculo das componentes *Wavelet* $\tilde{y}_{A_{m_0},t}$ e $\tilde{y}_{D_m,t}$, obtidos em (10) (ver, por exemplo, Haven *et al.* (2012)), preservando a autocorrelação em y_t e suas componentes, em (10), para todo t ($t = 1, \dots, T$). Com efeito, em uma decomposição *Wavelet* de nível 2 (isto é, $r = 2$), onde o truncamento, em (9), ocorre na segunda componente de detalhe, adota-se $m_0 = 2$, de modo que o estado y_t é expandido tal como: $y_t = y_{t_{A_2}} + y_{t_{D_2}} + y_{t_{D_3}} + \varepsilon_t$, em que ε_t é o erro de aproximação em t .

2.1.2 Modelos ARIMA (p, d, q)

Os *Modelos Autorregressivos Integrados de Médias Móveis* (ARIMA) são aplicáveis à modelagem de séries temporais estacionárias de segunda ordem (isto é, média e covariância constantes). De acordo com Hamilton (1994), se y_t ($t = 1, \dots, T$) é uma série temporal não estacionária na média, que pode se tornar estacionária, quando isso é possível, a partir de sua diferenciação, isto é, com a aplicação do operador diferença $\nabla^d y_t = W_t = (1-B)^d y_t$, onde: $\nabla^d := (1-B)^d$, sendo d um parâmetro inteiro positivo, e W_t ($t = d+1, \dots, T$) é a

versão da série temporal subjacente diferenciada d vezes. Neste caso, diz-se que W_t ($t = 1, \dots, T$) é um processo de séries temporais $ARMA(p, q)$ e é representado genericamente pelo modelo $\phi(B) W_t = \theta(B) \varepsilon_t$, onde: ε_t ($t = d + 1, \dots, T$) é uma realização de um processo estocástico independente e identicamente distribuído (i.i.d.); e p e q representam, respectivamente, a ordem dos polinômios autorregressivo $\phi(B)$ (AR(p)) e de médias móveis $\theta(B)$ (MA(q)), $\phi(B) := (1 - \phi_1 B - \dots - \phi_p B^p)$ e $\theta(B) := (1 - \theta_1 B - \dots - \theta_q B^q)$, sendo que ϕ_i ($i = 1, \dots, p$) e θ_i ($i = 1, \dots, q$) consistem em listas de números complexos que satisfazem às condições de estacionariedade e invertibilidade (BOX *et al.*, 2008).

Em outras palavras, pode-se dizer que o modelo ARIMA (p, d, q) é tal que combina linearmente valores passados das entradas W_t e choques aleatórios a_t descorrelacionados, de média zero e covariância constante, conforme em (11).

$$W_t = \alpha_1 W_{t-1} + \dots + \alpha_p W_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (11)$$

Ou, equivalentemente,

$$\phi(B) \nabla^d y_t = \theta(B) a_t \quad (12)$$

Um processo estacionário W_t , obtido por meio do operador diferença $\nabla^d y_t$ em que a série original y_t é não-estacionária, é chamado de não estacionário homogêneo (BOX *et al.*, 2008).

O modelo ARIMA (p, d, q) da equação (12) pode ser representado em termos de valores prévios de y_t e do valor atual e prévios de a_t , que é a forma usual do modelo e é útil para o cálculo de previsões conhecidas na forma de equações de diferenças:

$$y_t = \varphi_1 y_{t-1} + \dots + \varphi_{p+d} y_{t-p-d} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (13)$$

Em termos de modelagem, a análise do perfil dos gráficos das *Funções de Autocorrelação* (FAC) e das *Funções de Autocorreções Parciais* (FACP) pode indicar ordens p , d , q plausíveis para a obtenção de modelos iniciais (HAMILTON, 1994). Após a fase da identificação do modelo, passa-se à próxima etapa que é a estimação dos parâmetros, momento em que é necessário o uso de métodos iterativos não lineares de mínimos quadrados (geralmente, os de *Máxima Verossimilhança* (MV) e de *Mínimos Quadrados Ordinários* (MQO)), os quais se encontram detalhados em Box e Jenkins (1970) e Morettin e Toloi (2006).

Uma vez obtidas as estimativas dos parâmetros de um modelo ARIMA (p, d, q) , ocorre a sua validação que é por meio de testes de diagnósticos, para a verificação de pressupostos, como, por exemplo: teste de Box-Pierce, teste do periodograma acumulado, teste da autocorrelação cruzada, teste BDS, teste Arch-LM (teste de Multiplicador de Lagrange), estatística de Durbin-Watson, teste de Ljung-Box (FERNANDES, 2006).

Se y_t ($t = 1, \dots, T$) apresenta estrutura de sazonalidade, o modelo ARIMA adequado é apresentado, genericamente, em (14).

$$\phi(B) \Phi(B^S) \nabla_S^D \nabla^d y_t = \theta(B) \Theta(B^S) \varepsilon_t \quad (14)$$

onde: $\phi(B)$ e $\theta(B)$ são, respectivamente, os polinômios da parte autorregressiva e de médias móveis do modelo não sazonal; $\Phi(B^S) := (1 - \Phi_1 B^S - \dots - \Phi_p B^{pS})$ e $\Theta(B^S) := (1 - \Theta_1 B^S - \dots - \Theta_q B^{qS})$ são, respectivamente, a parte autorregressiva e de médias móveis da componente de sazonalidade da série temporal subjacente; $\nabla_S^D := (1 - B^S)^D$ é o operador de diferença sazonal de ordem D ; $\nabla^d := (1 - B)^d$ é o operador de diferença não sazonal de ordem d ; S é o período sazonal, $\phi_k \in \mathbb{C}$ e $\theta_j \in \mathbb{C}$ são os coeficientes complexos dos polinômios não sazonais e $\Phi_m \in \mathbb{C}$ e $\Theta_n \in \mathbb{C}$ são os coeficientes complexos dos polinômios sazonais (LUTKEPOHL, 2006).

2.1.3 Modelos ARIMAX-GARCH

Se y_t ($t = 1, \dots, T$) é uma série temporal estacionária (ou não estacionária que pode se tornar estacionária por diferenciação) que apresenta estrutura de autodependência linear (GUPTA e WILTON (1987)), assume-se então que $\left((x_{1,t})_{t=1}^T, \dots, (x_{(r+1),t})_{t=1}^T\right)$ é a sequência de $(r+1)$ vetores de realização de $(r+1)$ variáveis exógenas estacionária da série temporal y_t ($t = 1, \dots, T$). Com base em Box e Tiao (1975) e Pankratz (1991), cada realização y_t pode ser representada usando a modelagem ARIMAX (p, d, q) em que o modelo matemático genérico é formulado conforme segue abaixo:

$$\nabla^d y_t = \phi_0 + \sum_{j=1}^p \phi_j y_{t-j} + \sum_{i=1}^q \theta_i e_{t-i} + \sum_{i=1}^{r+1} \sum_{l_i=0}^{L_i} \alpha_{il_i} x_{i,t-l_i} + e_t \quad (15)$$

onde B é o operador de atraso definido por $B^k y_t := y_{t-k}$, com $k \in \mathbb{Z}$; $\nabla^d := (1 - B)^d$ é o operador diferença, com d representando sua ordem; $(\phi_i)_{i=0}^p, (\theta_j)_{j=1}^q$ são as listas ordenadas de parâmetros complexos, onde $\phi_p \neq 0$ e $\theta_q \neq 0$ e $(\alpha_{il_i})_{l_i=0}^{L_i}$ é a i -ésima lista ordenada de vetor de parâmetros complexo associado à componente exógena (estes três vetores de parâmetros precisam satisfazer tanto a invertibilidade quanto as condições de estacionariedade de segunda ordem) (HAMILTON, 1994; LUTKEPOHL, 2006); e_t é a inovação no tempo t que consiste a realização de uma variável aleatória ε_t em que ε_t ($t = 1, \dots, T$) é um processo estocástico com média zero e covariância constante; e p e q são, respectivamente, as ordens das partes autorregressivas (AR) $\sum_{j=1}^p \phi_j y_{t-j}$ (notação: AR(p)) e das partes médias móveis (MA) $\sum_{i=1}^q \theta_i e_{t-i}$ (notação: MA(q)). L_i é a maior ordem do lag em $\sum_{l_i=0}^{L_i} \alpha_{il_i} x_{i,t-l_i}$ da variável exógena $x_{i,t}$, onde $i=1, \dots, (r+1)$. Particularmente, se $\alpha_{il_i} = 0$ é para todo $i = 0, \dots, (r+1)$ e todo $l_i=0, \dots, L_i$ então o modelo (15) se torna o tradicional modelo ARIMA (p, d, q) . Por sua vez, um modelo estocástico SARIMAX $(p, d, q) \times (P, D, Q)_S$, também conhecido como modelo multiplicativo ARIMAX (p, d, q) ,

pode ser usado para modelagem de série temporal sazonal (HAMILTON, 1994) e consiste, genericamente, de um modelo ARIMAX com componentes sazonais.

Note que, em (15), o termo do erro e_t , para todo t , é a realização de uma variável aleatória não correlacionada ε_t com média zero e variância condicional constante, isto é, $\sigma_t^2 = \sigma^2$ (para todo t). Por outro lado, algumas séries temporais não satisfazem a condição de variância constante (homocedasticidade). Em alguns casos, a variação da variância condicional (volatilidade) pode depender de resíduos quadráticos do passado da série temporal ou da variância do passado de tal forma que a estrutura ARMA, bem como uma extensão da mesma, podem ser adotadas para temporalmente projetar a volatilidade. Desta forma, a variância incondicional é constante, mas se condicional a estes valores passados poderá variar com o tempo. Segundo Bollerslev (1986), o modelo geral GARCH (R,V) para a variância condicional σ_t^2 da inovação ε_t é dado por:

$$\sigma_t^2 = \delta_0 + \sum_{i=1}^R \delta_i e_t^2 + \sum_{j=1}^V \gamma_j \sigma_{(t-j)}^2 \quad (16)$$

com as seguintes restrições: $\sum_{i=1}^R \delta_i + \sum_{j=1}^V \gamma_j < 1$; $\gamma_0 > 0$; e $\gamma_j \geq 0$ e $\delta_i \geq 0$ ($i = 1, \dots, R$) e ($j=1, \dots, V$). Essa função é usada pelo método WARIMAX-GARCH Neural para gerar previsões dentro e fora da amostra para a variância condicional da série temporal y_t ($t = 1, \dots, T$). Além disso, σ_t^2 multiplicado pelo parâmetro φ pode ser usado em (15) de tal forma que este efeito não linear é utilizado no cálculo da previsão média. Define-se o mecanismo gerador de um processo de inovação e_t ($t = 1, \dots, T$) por $e_t := \sigma_t z_t$, onde: z_t é a realização de uma variável aleatória padronizada Z_t , independente e identicamente distribuída e σ_t é o desvio padrão condicional proveniente de e_t .

Um modelo composto por (15) e (16), que apresenta as restrições acima mencionadas, é chamado de modelo ARIMAX-GARCH. Então, para obter o melhor modelo ARIMAX-GARCH possível, três passos básicos devem ser

realizados: (i) testar os valores plausíveis para os parâmetros p, d, q, i e L em (15) bem como os parâmetros V e R , em (16) (os quais podem ser obtidos por meio da análise do perfil dos gráficos da autocorrelação e autocorrelação parcial dos resíduos ordinários z_t ($t = 1, \dots, T$) e resíduos quadráticos z_t^2 ($t = 1, \dots, T$), como descrito por Bollerslev (1986) e Engle (1982)) (ii) definir o método a ser utilizado para a estimação dos parâmetros ARIMAX-GARCH (o mais comum é o método de estimação da Máxima Verossimilhança – MLE – (HAMILTON, 1994)); (iii) fazer a verificação de diagnóstico para escolher o modelo parcimonioso e adequado a ser usado para gerar previsões dentro e fora da amostra de y_t ($t = 1, \dots, T$) e da sua volatilidade.

2.1.4 Redes Neurais Artificiais

Uma *Rede Neural Artificial* (RNA) é uma das técnicas de Inteligência artificial que consiste de modelos formados por unidades de processamentos simples (neurônios artificiais) baseados no sistema nervoso central biológico (Haykin, 2001). Uma RNA realiza operações similares ao cérebro humano por meio de conhecimento experimental alcançado previamente a partir de padrões tais como: *aprendizado, associação, generalização e abstração*. Uma RNA com uma, e somente uma, camada escondida consiste de sistemas formados por tais neurônios artificiais (HAYKIN, 2001; ABELÉM, 1994) e conforme Cybenko (1989), pode ser utilizada para aproximar classes de *funções contínuas de suporte compacto* (KUBRUSLY, 2001). A disposição dos neurônios artificiais em uma RNA pode ser formada por camadas neurais interligadas por um amplo número de conexões, chamados pesos sinápticos, que são ponderações dos padrões de entrada recebidos (RUSSELL; NORVIG, 1995; TAFNER, 1996).

A esquematização simplificada de um neurônio biológico, em cuja estrutura um neurônio artificial de uma RNA é inspirado, é mostrada na Figura 5.

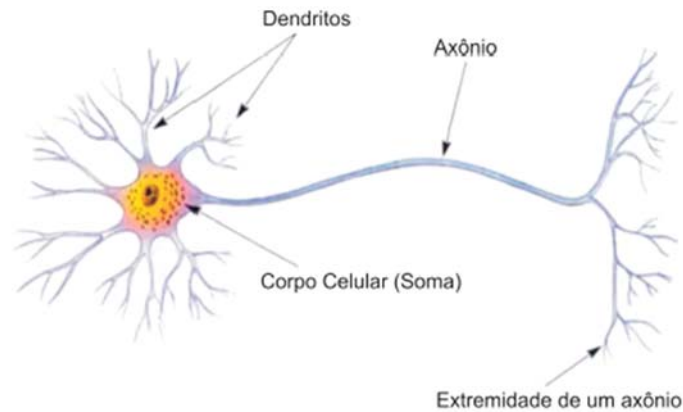


Figura 5 - Esquematização simplificada de um neurônio biológico

A constituição de um neurônio biológico, segundo Tafner (1996) é dada basicamente, pelos seguintes elementos: um *corpo celular* (ou *soma*), que contém o núcleo da célula; diversos *dendritos*, através dos quais impulsos elétricos são recebidos; e um *axônio*, por meio do qual são enviados impulsos elétricos. As interligações entre neurônios são efetuadas através de *sinapses* (isto é, pontos de contato que são controlados por impulsos elétricos e por reações químicas) entre os dendritos e os axônios, formando, assim, uma rede neural biológica de transmissão de informações. Segundo Haykin (2001), um neurônio artificial é constituído pelos elementos descritos, em (1), (2), (3) e ilustrados, na Figura 6.

- (1) Um *conjunto de pesos sinápticos*, denotado por $\{\{W_{kt}\}_{k=1}^K\}_{t=1}^T$, e um *conjunto de padrões de entrada*, denotado por $\{x_t\}_{t=1}^T$, onde W_{kt} denota o peso sináptico associado ao padrão de entrada x_t (observado no instante t) e ao k -ésimo neurônio artificial. Cada padrão de entrada x_t é multiplicado pelo peso respectivo sináptico W_{kt} , ou seja, $x_t W_{kt}$, para todo $t \in \{1, \dots, T\}$ e para todo $k \in \{1, \dots, K\}$;
- (2) Um *somador* que consiste de um operador denotado por \sum , para realizar a operação de soma dos padrões de entrada em $\{x_t\}_{t=1}^T$ ponderados (por meio da operação de multiplicação) pelos respectivos pesos sinápticos no conjunto $\{\{W_{kt}\}_{k=1}^K\}_{t=1}^T$. Isto é, $\sum_{t=1}^n x_t W_{kt} + b_k$, onde: x_t é o padrão de entrada, em t , W_{kt} é o peso sináptico associado

ao padrão de entrada x_t e ao k -ésimo neurônio artificial; e b_k é o intercepto (ou nível, ou *bias*), que tem a função de aumentar ou diminuir a entrada líquida da função de ativação, podendo, desse modo, ser positivo ou negativo. A soma ponderada $\sum_{t=1}^n x_t W_{kt} + b_k$, de acordo com Haykin (2001), pode ser chamada de *regra de propagação*; e

- (3) Uma *função de ativação*, denotada por φ , que é utilizada para restringir a amplitude da saída y_k de um neurônio artificial, dada a soma ponderada $\sum_{t=1}^n x_t W_{kt} + b_k = v_k$. Isto é, $y_k = \varphi(v_k)$, onde $k \in \{1, \dots, K\}$. Além disso, segundo Herrera e Lozano (1998), são utilizadas na identificação de relações lineares ou não lineares entre os padrões de entrada e os de saída em modelos neuronais.

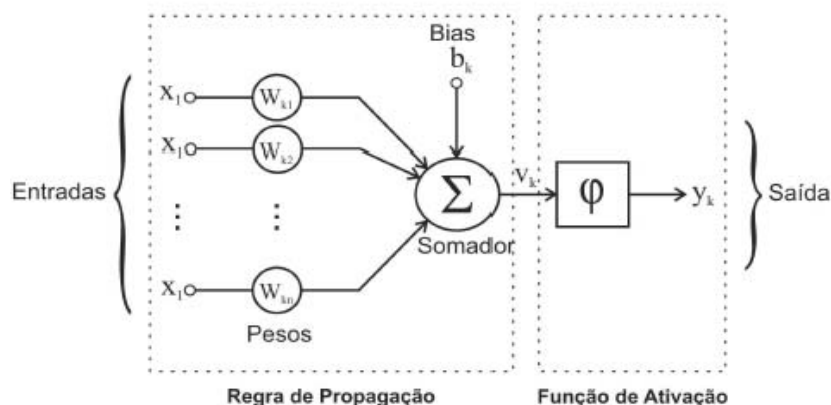


Figura 6 - Arquitetura básica de um neurônio artificial

A Rede Neural Artificial mais utilizada para previsões de séries temporais é a *Perceptron* com múltiplas camadas alimentadas para frente (*Feedforward Multilayer Perceptron*) (HAYKIN, 2001). Particularmente, uma RNA com uma camada escondida, é constituída por três camadas (camada de entrada, camada escondida e camada de saída) com unidades de processamento numérico (neurônios) conectadas por ligações acíclicas. A relação entre a saída y_t e as entradas defasadas y_{t-k} ($k = 1, \dots, L$) tem representação matemática dada por

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g(\beta_{0j} + \sum_{i=1}^L \beta_{ij} y_{t-i}) + \varepsilon_t \quad (17)$$

onde α_j ($j = 0, 1, \dots, q$) e β_{ij} ($i = 0, 1, \dots, L; j = 0, 1, \dots, q$) são os parâmetros da RNA, chamados de pesos sinápticos; L é o número de nós de entrada, q é o número de nós da camada escondida; ε_t é o erro de aproximação no tempo t e $g(\cdot)$ é a função transferência (HAYKIN, 2001).

2.1.4.1 Componentes Básicas de uma RNA

De acordo com Zsolt (2006), uma RNA pode ser classificada conforme as características: (a) *função de ativação* dos neurônios artificiais [seção 2.1.4.1.1]; (b) a *topologia* ou *arquitetura*, que consiste na forma como os neurônios estão interligados na estrutural neuronal [seção 2.1.4.1.2]; e (c) o *treinamento neural* [seção 2.1.4.1.3].

2.1.4.1.1 Função de Ativação

As funções de ativação mais empregadas, de acordo com Campos (2010), em aplicações que envolvem a projeção de séries temporais, são: *linear*, *degrau*, *logística*, *tangente hiperbólica*.

(a) Função Linear:

$$\sum_{t=1}^n x_t W_{kt} + b_k = v_k \quad (18)$$

De acordo com Kubrusly (2001), a *função linear*, comumente, é empregada nos neurônios da camada de saída e é indicada quando as sequências de padrões de saída associados a estes não são *conjuntos limitados* (TEIXEIRA JUNIOR, 2013).

(b) Função Degrau (ou Limiar):

$$y_k = \begin{cases} 1, & \text{se } v_k \geq 0 \\ 0, & \text{se } v_k < 0 \end{cases} \quad (19)$$

É a função utilizada para valores binários e o conjunto imagem (saída do neurônio) da *função degrau* (ou *limiar*) assume valores 0 ou 1. Conforme

Haykin (2001), a função degrau pode ser chamada particularmente de *função de sinal* quando a sua imagem assumir valores -1 ou +1.

(c) **Função Sigmoid Logística:**

$$y_k = \frac{1}{1 + \exp\{-\alpha v_k\}}, \text{ onde } \alpha \in \mathbb{Z} \quad (20)$$

Segundo Kubrusly (2001), a *função sigmoide logística* admite valores no *intervalo limitado e fechado* $[0,1]$ e consiste em uma *transformação monotonamente crescente* (LIMA, 2006), que permite a transição gradual entre as funções linear e degrau, onde a regra de propagação do k-ésimo neurônio artificial é dada por $\sum_{t=1}^n x_t W_{kt} + b_k = v_k$.

(d) **Tangente Hiperbólica:**

$$y_k = \tanh(\alpha v_k), \text{ onde } \alpha \in \mathbb{Z} \quad (21)$$

A *função tangente hiperbólica*, denotada por $\tanh(.)$, é a função sigmoide cuja saída de um neurônio artificial assume valores no intervalo fechado $[-1,1]$.

2.1.4.1.2 Topologia

A *topologia* ou *arquitetura* de uma RNA é uma importante característica que se refere à maneira como os seus neurônios artificiais estão interconectados, baseada nas interconexões do cérebro humano, de forma que as informações possam ser processadas de modo dinâmico ou interativo. Quanto à arquitetura, Russell e Norvig (1995), determinam dois tipos básicos de RNA's, que são a *feedforward* (*não recorrentes* ou *diretas*) e a *recorrente*, conforme itens (a) e (b), respectivamente:

- (a) **RNA's com topologia *feedforward* (diretas):** as RNA's com topologia *feedforward* não apresentam conexões entre neurônios localizados na mesma camada ou entre um neurônio de uma camada com outro neurônio de uma camada anterior (os grafos não têm ciclos). Na Figura 7 (adiante, página 48), é ilustrada uma RNA *feedforward*, que contém

uma camada escondida e um neurônio na camada de saída. A composição de uma RNA *feedforward*, de acordo com Haykin (2001), é apresentada basicamente por um conjunto de nós de entradas, os quais fazem a distribuição dos padrões de entrada para os neurônios artificiais; por uma camada de saída, cujos neurônios enviam os resultados do processamento da rede neural; e por camadas intermediárias, cujas saídas dos neurônios são as entradas dos neurônios da camada seguinte, é a camada onde ocorrem os processamentos internos da rede (HAYKIN, 2001).

- (b) **RNA's, com topologia *feedback* (recorrentes):** as RNA's com arquitetura *feedback* são aquelas que possuem realimentação das saídas para as entradas (o grafo de conectividade contém pelo menos um ciclo), de modo que as saídas destas são determinadas pelas entradas atuais e pelas saídas anteriores (HAYKIN, 2001). A topologia das RNA's *feedback* pode apresentar interconexões entre neurônios da mesma camada ou entre neurônios de camadas anteriores.

A topologia de uma RNA (básica) (RUSSELL; NORVIG, 1995) envolve, geralmente, os parâmetros: padrões de entrada, número de camadas intermediárias, quantidade de neurônios por camada (entrada, intermediária e de saída), tipo de função de ativação dos neurônios e os padrões de saída.

2.1.4.1.3 Treinamento Neural

O treinamento de uma rede neural artificial, segundo Abelém (1994), consiste de um método de otimização, a partir do uso de um algoritmo de otimização, em que é efetivado o ajuste numérico dos pesos sinápticos que visam a maximização ou minimização de uma função objetivo (RAGSDALE, 2004). Maiores detalhes sobre os algoritmos de otimização são encontrados em Vanderplaats (1999).

O treinamento de uma RNA, geralmente, é dado a partir da minimização da função *Erro Quadrático Médio* (EQM) (Box *et al.*, 2008) entre os padrões gerados na saída da RNA (resultados) e os valores reais (desejados).

Comumente, segundo Haykin (2001), se utiliza um único conjunto de sinais como padrões de entrada e de saída, respectivamente, denominado amostra de treino (CAMPOS; 2010).

O uso de padrões de entrada e de saída deve ser representativo, segundo Herrera e Lozano (1998), quanto a uma população de interesse (MORETTIN e TOLOI, 2006), além disso, estes padrões devem conter uma quantidade de sinais suficientemente grande para uma modelagem apropriada. Os procedimentos de treinamento neural, segundo Haykin (2001), ocorrem em dois tipos básicos:

- (a) **Treinamento supervisionado:** os padrões de treinamento são compostos pelos conjuntos de entrada e de saída desejadas que são apresentados à rede. As entradas são apresentadas à rede neural artificial, que realiza seu processamento (treinamento) e, em seguida, cada saída neuronal é comparada com a respectiva saída esperada; calcula-se, assim, um sinal de erro (diferença entre as duas saídas). A cada ciclo, o algoritmo de treinamento faz a atualização dos pesos sinápticos com base nos valores dos erros obtidos nos padrões de treino, de modo que as saídas da RNA se aproximem o máximo possível das saídas desejadas (erro mínimo ou aceitável). O algoritmo de aprendizado supervisionado mais comum é o *Backpropagation* (HAYKIN, 2001).
- (b) **Treinamento não supervisionado:** é caracterizado pela ausência de algum elemento externo supervisor, ou seja, os padrões de treinamento não possuem padrões de saída desejados; nesse caso não existem comparações que permitam a determinação de sinais de erro. O processo de treinamento não supervisionado ou auto-organizado, em geral, captura propriedades estatísticas do conjunto de padrões de entrada, segundo Palit e Popovic (2005), e são formados agrupamentos disjuntos de padrões similares, com base em uma topologia. Ainda de acordo com Palit e Popovic (2005), o

processo de treinamento não supervisionado, em geral, extrai as propriedades estatísticas do conjunto de padrões de entrada, de maneira a formar agrupamentos disjuntos constituídos de padrões similares, com base em uma *topologia* (KUBRUSLY, 2001). Os algoritmos de aprendizado não supervisionado mais importantes na literatura são: Algoritmo de Hopfield e Mapas de Kohonen (HAYKIN, 2001).

Para as redes neurais artificiais com treinamento supervisionado (que são de interesse nesta pesquisa), o erro quadrático médio é o critério mais utilizado (nas amostras de treino, validação e teste) para a sua escolha.

2.1.4.1.4 Normalização dos Padrões

De acordo com Tafner (1996), a *normalização dos padrões* de entrada e de saída de RNA's, com treinamento supervisionado, é fundamental para que suas magnitudes fiquem dentro de uma faixa de valores, a fim de evitar que os padrões com maiores magnitudes distorçam os valores dos pesos sinápticos no processo de treinamento neural e de suprimir problemas associados à *paralisia neural* (HAYKIN, 2001). Neste caso, os padrões devem ser normalizados em um intervalo limitado compatível com o domínio das funções de ativação dos respectivos neurônios da estrutura neuronal. De acordo com Campos (2010), as normalizações de sinais mais comuns, em aplicações envolvendo projeções de séries temporais, são as que as transformam em pontos nos intervalos limitados e fechados $[0,1]$ ou $[-1,1]$; ou, ainda, em pontos *padronizados* (ou seja, elementos de um conjunto de sinais cuja média é igual a 0 e desvio padrão, igual a 1).

Nesta perspectiva, as normalizações de padrões *MINIMAX*, *método premnmx*, *z-escore*, *sigmoidal* são, respectivamente, definidas nos itens (a), (b), (c) e (d), abaixo. Para tanto, assumindo previamente que $\{x_t\}_{t=1}^n$ é, simultaneamente, o conjunto dos padrões de entrada e de saída de uma RNA, com treinamento supervisionado, e que $\{x_t^m\}_{t=1}^n$ é a sua versão normalizada.

- (a) **Método MINIMAX:** utiliza o valor máximo do conjunto $\{abs(x_t)\}_{t=1}^n$ de sinais, o qual é denotado por $Max\{abs(x_t)\}_{t=1}^n$, onde $abs(x_t)$ denota o valor absoluto de x_t . A normalização dos pontos da sequência $\{x_t\}_{t=1}^n$ ocorre conforme a transformação definida em (22), que toma valores no intervalo $[-1,1]$.

$$x_t^m = x_t / Max\{abs(x_t)\}_{t=1}^n \quad (22)$$

- (b) **Método premnmx:** utiliza os valores máximo e mínimo do conjunto de padrões $\{x_t\}_{t=1}^n$, os quais são denotados, respectivamente, por $Max\{x_t\}_{t=1}^n$ e $Min\{x_t\}_{t=1}^n$. A normalização de cada ponto x_t ocorre de acordo com a transformação definida em (23), que toma valores no intervalo $[-1,1]$.

$$x_t^m = 2 \times \left[\left(x_t - Min\{x_t\}_{t=1}^n \right) / \left(Max\{x_t\}_{t=1}^n - Min\{x_t\}_{t=1}^n \right) \right] - 1 \quad (23)$$

- (c) **Método do z-escore:** realiza a normalização do conjunto $\{x_t\}_{t=1}^n$ em torno de sua média, denotada por \bar{x}_n , utilizando o seu desvio padrão, denotado por S_x , de acordo com a transformação definida em (24). A média de conjunto $\{x_t^m\}_{t=1}^n$ de padrões normalizados é igual a 0 e o desvio padrão, igual a 1.

$$x_t^m = (x_t - \bar{x}_n) / S_x \quad (24)$$

- (d) **Método Sigmoidal:** utiliza a média e o desvio padrão do conjunto $\{x_t\}_{t=1}^n$ de padrões, os quais são, respectivamente, denotados por \bar{x}_n e S_x . A normalização de cada ponto x_t em $\{x_t\}_{t=1}^n$ ocorre de acordo com a transformação definida em (25), que toma valores no intervalo $[0,1]$.

$$x_i^m = \frac{1}{1 + \exp\{(x_i - \bar{x}_n)/S_x\}} \quad (25)$$

2.1.4.2 RNA's multicamadas *Feedforward*, com uma Camada Escondida

Segundo Haykin (2001), uma propriedade importante de uma RNA é a sua capacidade de processar sinais de entrada não lineares. Uma RNA com múltiplas camadas tem a presença de camada (s) oculta(s), cujos nós são chamados de neurônios ocultos, e tem como função a intervenção entre a entrada externa e a saída da rede de maneira útil. Haykin (2001) afirma, ainda, que ao serem adicionadas uma ou mais camadas ocultas à rede, se torna capaz de extrair estatísticas de ordem elevada. A construção de uma rede multicamada leva em consideração a determinação do número de camadas escondidas, o número de neurônios por camadas escondidas e a especificação dos pesos sinápticos que fazem a interconexão entre os neurônios nas camadas distintas da rede.

De acordo com Rezende (2003), a camada de saída final da rede é responsável pelo recebimento dos estímulos da camada intermediária e, também, pela construção da camada global para o padrão de ativação fornecido pelos nós de fonte de camada de entrada. As camadas intermediárias são responsáveis pela coleta das características dos padrões de entrada e inserção dos pesos para essas características de forma que a rede crie uma representação mais precisa e complexa do problema.

A seguir, conforme Figura 7, têm-se: na parte (a), a ilustração de uma arquitetura de uma RNA multicamada (MLP) *feedforward*, com uma camada escondida com 4 neurônios, 1 neurônio na camada de saída e janela J igual a 6; e, na parte (b), as duas fases do *algoritmo de treinamento backpropagation* (HAYKIN, 2001; RUSSELL, NORVIG 1995): a fase *forward* e a fase *backward*.

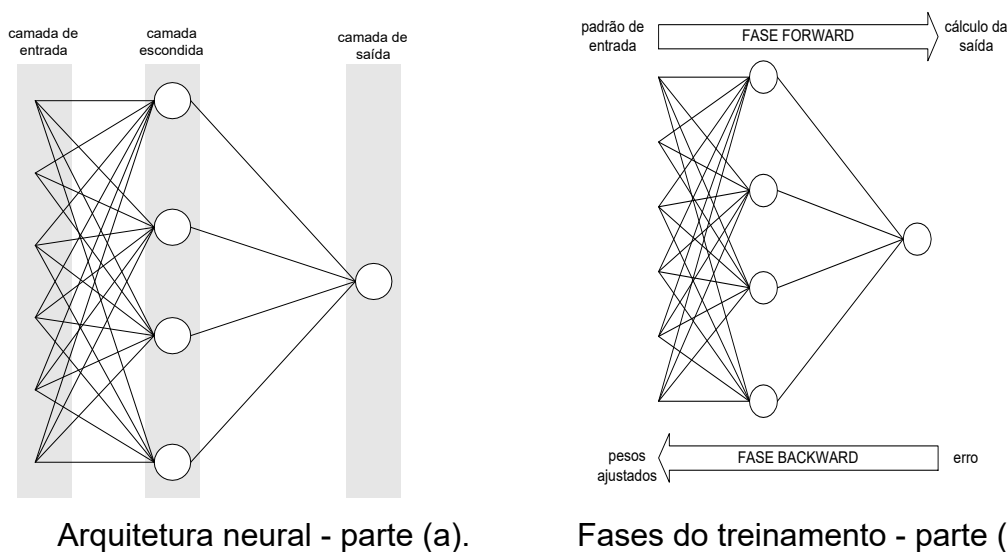


Figura 7 - RNA MLP *feedforward*

Um dos principais algoritmos de treinamento neural é o *backpropagation*. O ajustamento numérico dos pesos sinápticos, com o uso do algoritmo *backpropagation*, é realizado por um processo de otimização constituído de duas fases: a *forward* e a *backward*. Na fase *forward*, a partir de um conjunto de padrões de entrada é calculada a resposta fornecida pela RNA, que varia de acordo com tamanho o da janela neural atribuída. Na fase *backward*, é feita a atualização dos pesos sinápticos com base no erro entre a resposta da RNA (prevista) e a resposta desejada (real). No decorrer da etapa de treinamento, são apresentados à RNA os padrões de entrada e as respectivas respostas desejadas, de modo a minimizar o erro quadrático médio, da amostra de treino, entre os sinais de saída da RNA e os sinais de saída desejados.

2.1.4.3 Geração de Previsões via Redes Neurais Artificiais

É possível prever valores futuros de uma série temporal com o uso de uma RNA *feedforward*, com treinamento supervisionado, com uma camada escondida e uma camada de saída com um neurônio artificial uma vez que suas observações (sinais) podem ser interpretadas como pontos de uma função de suporte compacto e, conforme Teixeira Júnior (2013), estas RNA's

são aproximadores deste tipo de mapeamento na topologia $\| \cdot \|_{\infty}$. O tamanho da janela neural L e do horizonte de previsão h desejado é definido previamente. Em uma série temporal com estrutura de autodependência autorregressiva, linear ou quadrática (HAMILTON, 1994), os valores passados da série temporal que se deseja projetar formam o conjunto de padrões de entrada enquanto os valores a serem previstos compõem o padrão de saída.

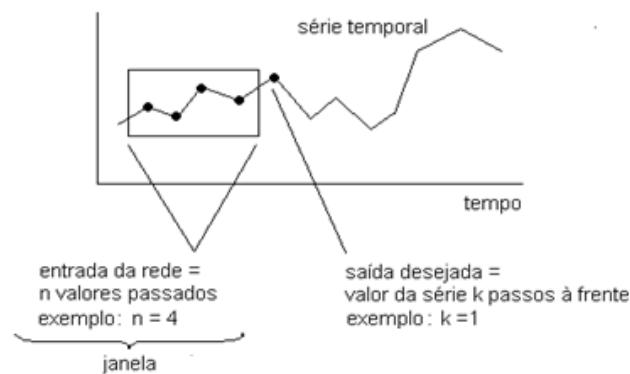


Figura 8 - Ilustração da dinâmica de previsão, um passo à frente, de uma série temporal, com estrutura de autodependência igual a 4, através de uma RNA

A Figura 8 ilustra a previsão, com horizonte de previsão $h = 1$ (ou um passo à frente), de uma série temporal com a presença de estrutura de autodependência cuja defasagem é igual a 4 (ou seja, $L = 4$), através de uma RNA *feedforward*. O vetor de padrão de entrada é formado pelos últimos 4 sinais temporais que antecedem o padrão de saída (ou valor a ser previsto). De acordo com Campos (2010), para prever o sinal subsequente, a janela da RNA é deslocada um passo à frente.

2.1.5 Combinação de Previsões

Os pioneiros na área de combinações de previsões foram Bates e Granger (1969), que sugeriam uma combinação, representada em (26), para a previsão da série temporal y_t com $(t=1,2,...,T)$:

$$\hat{y}_t = \rho_1 \times \hat{y}_{1,t} + \rho_2 \times \hat{y}_{2,t} \quad (26)$$

Note que em (26), \hat{y}_t é a previsão combinada linearmente enquanto as previsões realizadas através de dois métodos, aqui designados por método 1 e método 2, são $\hat{y}_{1,t}$ e $\hat{y}_{2,t}$, respectivamente. Os coeficientes (pesos) ρ_1 e ρ_2 são obtidos a partir da matriz de covariância de erros de previsões individuais.

Granger e Ramathan (1984) propuseram uma combinação linear sem a restrição para os coeficientes ($\rho_1 + \rho_2 = 1$) e com adição de uma constante que tornava a previsão não viesada. Em 1987, Gupta e Wilton (1987) propuseram a utilização da média aritmética simples das previsões dos modelos individuais. Teixeira Junior (2013) empregou a programação matemática para obter os valores dos coeficientes ρ_1 e ρ_2 a partir da resolução de um problema de programação não linear, cuja função objetivo era composta pela soma dos erros quadráticos de previsão.

A combinação de previsões permite aumentar a acurácia das previsões, de acordo com Granger (1989). Hollauer *et al.*, (2008), salientam que a diversificação de previsões combinadas leva à diminuição do erro, enquanto Faria e Mubwandarikwa (2008) afirmam que a previsão combinada agrega informações peculiares oriundas dos diferentes métodos preditivos em relação à flutuação estocástica da série temporal a ser modelada. Wallis (2011) aborda uma revisão histórica com muitos dos principais artigos relacionados à combinação de métodos preditivos individuais, revisados em um período de 42 anos, os quais comprovam sua eficiência, além de abordarem uma ampla diversidade de aplicações.

Considere o caso particular em que, a partir da modelagem das componentes *Wavelets* que resulta em $p + 1$ previsões obtidas por ARIMA-GARCH e $p + 1$ obtidas por RNA, a técnica da combinação consiste em fazer $p + 1$ combinações lineares em que cada uma envolve os pares de estimativas $(\hat{A}_{p,RNA} e \hat{A}_{p,ARIMA})$, $(\hat{D}_{1,RNA} e \hat{D}_{1,ARIMA})$, ..., $(\hat{D}_{p,RNA} e \hat{D}_{p,ARIMA})$. A combinação linear para a componente da aproximação $(\hat{A}_{p,CL})$ está descrita em (27),

$$\hat{A}_{p,CL} = \alpha_{Ap,ARIMA} \times \hat{A}_{p,ARIMA} + \alpha_{Ap,RNA} \times \hat{A}_{p,RNA} \quad (27)$$

em que $\alpha_{Ap,ARIMA}$ e $\alpha_{Ap,RNA}$ são parâmetros a serem obtidos por meio da otimização do problema de programação não linear (28), cuja função objetivo é a raiz do erro quadrático médio (RMSE).

$$\text{Min } RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{A}_{p,CL} - A_p)^2} \quad (28)$$

sujeito a $\alpha_{Ap,ARIMA}$ e $\alpha_{Ap,RNA}$ irrestritas

Analogamente, as combinações lineares para as componentes de detalhes $\hat{D}_{1,CL}$ (29)... $\hat{D}_{p,CL}$ (30) são dadas por

$$\hat{D}_{1,CL} = \alpha_{D1,ARIMA} \times \hat{D}_{1,ARIMA} + \alpha_{D1,RNA} \times \hat{D}_{1,RNA} \quad (29)$$

⋮

$$\hat{D}_{p,CL} = \alpha_{Dp,ARIMA} \times \hat{D}_{p,ARIMA} + \alpha_{Dp,RNA} \times \hat{D}_{p,RNA} \quad (30)$$

A previsão $\{\hat{y}_t\}_{t=1}^T$ da série observada $\{y_t\}_{t=1}^T$ poderá ser obtida por combinação linear, equação (31),

$$\hat{y} = \alpha_{A,CL} \times \hat{A}_{p,CL} + \alpha_{D1,CL} \times \hat{D}_{1,CL} + \dots + \alpha_{Dp,CL} \times \hat{D}_{p,CL} \quad (31)$$

sendo os parâmetros $\alpha_{A,CL}, \alpha_{D1,CL}, \dots, \alpha_{Dp,CL}$ calculados a partir de problema similar a (28), com as devidas trocas de variáveis.

2.2 Revisão da Literatura

A análise e a previsão de séries temporais em diversas áreas, como ciências, engenharias e negócios, têm sido bastante exploradas e são de grande interesse aos cientistas e engenheiros. Morettin e Tolo (2006) definem um processo estocástico como uma família de variáveis aleatórias do tipo $Z(t) = Z_t$, $t \in T$ tal que para cada $t \in T$, $Z(t)$ é uma variável aleatória. E, série temporal é o processo estocástico formado por um conjunto de observações ordenadas no tempo e que apresentam dependência serial entre elas.

Desde 1927 quando George Udny Yule lançou a ideia de que uma série temporal pode ser entendida como sendo gerada por um conjunto de variáveis aleatórias conjuntamente distribuídas, muitos foram os progressos na melhoria dos modelos com melhor acurácia em séries temporais, particularmente aqueles conhecidos por autorregressivos AR(p). Neste contexto Slutzky (1937) introduziu os modelos médias móveis MA(q) em 1937 e, em 1970, Box e Jenkins apresentaram um método que fez a integração dos conhecimentos existentes na época. Esse método consiste em ajustar modelos autorregressivos integrados com médias móveis (ARIMA) a um conjunto de dados a partir de três estágios: identificação do modelo, estimação dos parâmetros e verificação do modelo ajustado.

Para modelagens de séries temporais que possuem estrutura de dependência linear, ou seja, aquelas provenientes de processos estocásticos lineares, a metodologia de Box e Jenkins é amplamente utilizada. Já para as séries de tempo que possuem estrutura de dependência não linear e que são provenientes de processos não lineares, predominantemente, é indicado o uso de Redes Neurais Artificiais (RNA). Vale salientar que o uso de RNA serve tanto na abordagem linear como na não linear. Hsu e Chen (2003), Kurt e Oktay (2010), Li *et al.* (2014), Chang *et al.* (2014), Mattar *et al.* (2015), entre outros, utilizaram Redes Neurais Artificiais (RNA) na previsão de séries temporais.

Importantes avanços foram feitos em muitas áreas relacionadas à previsão de séries temporais. De Gooijer e Hyndman (2006) revisaram 25 anos de pesquisas de obras altamente influentes e apontam a obra de Box e Jenkins (1970) como sendo impactante na teoria e na prática da análise e previsão de séries temporais. O advento do computador, que tornou a computação veloz e barata, popularizou o uso dos modelos ARIMA (p, d, q). Outro fator importante no desenvolvimento da previsão foi o surgimento do Algoritmo de Marquardt em 1963, que possibilitou a obtenção de estimativas adequadas dos parâmetros do modelo (MARQUARDT, 1963). Além disso, são destacados os modelos ARCH/GARCH que fornecem previsão mais precisa em séries de

tempo que apresentem volatilidade implícita em diversos trabalhos. Não ficam de fora as redes neurais artificiais (RNA's) que desempenharam grande sucesso na previsão de séries temporais e que, apesar disso, ainda necessitam de um maior desenvolvimento. Destaca-se, também, a importância da estratégia de combinação de métodos lineares e não lineares, os quais apresentam, na maioria das vezes, resultados superiores em relação aos métodos isolados. Por fim, destaca-se a importância da determinação dos intervalos de confiança para as previsões pontuais.

Uma técnica bastante utilizada e que apresenta eficácia empiricamente comprovada é a da combinação de previsões. Bates e Granger (1969) foram um dos primeiros a estruturar um método de combinação de resultados de modelos de previsão em que, através da combinação linear, eram calculados os pesos a partir da matriz de covariância de erros das previsões individuais de cada modelo. Outros autores como Cavaleri e Ribeiro (2011), Fiordaliso (1998), Zou e Yang (2004), Bunn (1988), Martins (2011) desenvolveram trabalhos que comprovam a eficácia do método de combinar resultados previstos por modelos tradicionais. Clemen (1989) revisou diversos trabalhos da literatura e chegou à conclusão de que a precisão das previsões pode ser significativamente melhorada por meio da combinação de várias previsões individuais.

Uma técnica eficaz que pode ser utilizada no subsídio dos modelos de previsão de séries temporais é a decomposição *Wavelet*. Y. Mayer utilizou os resultados de Mallat para construir a primeira *Wavelet* não trivial (suave). Ingrid Daubechies utilizou os trabalhos de Mallat para a criação de uma família de bases ortonormais de *Wavelet* suaves com suportes compactos, cujo trabalho passou a ser uma das principais referências para as atuais aplicações da transformada *Wavelet* (LIMA, 2002).

Aplicações da decomposição *Wavelet* são encontradas em diversos trabalhos de previsão de séries temporais. Lima (2002) utilizou *Wavelet* para a separação da parte determinística dos ruídos em séries temporais econômicas.

Teixeira Júnior *et al.* (2012) fizeram a combinação de componentes *Wavelets* de uma série temporal para utilizar como padrões de entrada de uma RNA *FeedForward* MLP e cuja saída fornece a previsão da série temporal. Karthikeyan e Kumar (2013); Tiwari e Chatterjee (2010); Kisi (2010); Kisi e Cimen (2011); Nalley *et al.* (2012) também empregaram *Wavelet* em seus trabalhos de previsão de séries temporais.

Como exemplo de séries temporais voláteis, as séries temporais financeiras apresentam a volatilidade não constante no decorrer do tempo e esta influencia, por exemplo, nas tomadas de decisão de investidores sobre quando e quanto investir. Tais séries diferem em sua estimação e previsão por se fazer necessário um tratamento diferenciado daqueles obtidos nos modelos clássicos ARIMA, já que os modelos ARMA/ARIMA modelam apenas o nível da série e não a variância. É plausível estimar a volatilidade por meio da variância condicionada a partir de funções lineares e não lineares. O modelo ARCH (modelo autorregressivo de heteroscedasticidade condicional) foi introduzido por Engle (1982). O modelo ARCH generalizado denominado GARCH (modelo autorregressivo de heteroscedasticidade condicional generalizado), proposto por Bollerslev (1986) é uma extensão do modelo ARCH. Neste modelo considera-se que a variância condicionada do processo de erro está relacionada, além dos quadrados dos valores passados da série, com as variâncias condicionadas passadas. Liu e Shi (2013) mostraram que os modelos ARMA-GARCH-M são eficazes para modelar e prever a média e a volatilidade dos preços da eletricidade. Gabriel (2012) avaliou a capacidade de previsão dos modelos GARCH, dentro e fora da amostra, no mercado de ações romeno com resultados importantes no cálculo do valor de risco em processos de gestão de riscos. Byun e Cho (2013) abordaram o uso dos modelos GARCH para a previsão de preços futuros do carbono, e de acordo com os resultados, mostram que os preços do petróleo, carvão e eletricidade podem ser usados para a previsão da volatilidade do mercado futuro de carbono.

A incerteza na previsão de séries de tempo pode ser quantificada ao se utilizar o conceito de intervalos de confiança. Diversos trabalhos da literatura

utilizam alguma técnica para a construção desses intervalos. Charytoniuk e Niebrzydowski (1998), Clements e Kim (2007), Chlapinski e Rózanski (2013), O'Connor e Lawrence (1989) e Misiorek *et al.* (2006) utilizaram intervalos de confiança a partir de erros de precisão produzidos pelos métodos aplicados.

O avanço das tecnologias da informação e a elaboração de modelos matemáticos aperfeiçoados permitem fazer uso de previsões, as quais são empregadas por diversos profissionais de engenharias, ciências e negócios. Em diversas situações, o processo de tomadas de decisão que emprega previsões de séries temporais, a fim de analisar e planejar necessidades operacionais futuras se dá em diversas áreas, sendo citadas algumas delas – econômica: como índices de bolsa de valores, variação dos juros; planejamento energético: como produção e consumo de energia; fenômenos naturais: como radiação solar, precipitação atmosférica, entre outros.

3 MATERIAIS E MÉTODOS

Neste capítulo, mais especificamente na seção 3.1, serão apresentadas as séries temporais modeladas por meio da metodologia proposta WARIMAX-GARCH Neural e via modelagem ARIMA-GARCH. Além disso, na seção 3.2 serão descritas as etapas do método proposto, que são: decomposição ortogonal *Wavelet* de nível $r = 2$ da série temporal subjacente, modelagem e previsão de suas componentes *Wavelet* via modelo ARIMA-GARCH e via Redes Neurais Artificiais, modelagem e previsão, via modelo ARIMAX, da série subjacente com entrada (input) das variáveis artificiais exógenas *Wavelet* completadas, criação dos intervalos de previsão.

3.1 Materiais

Os diversos instrumentos da usina possuem natureza, tamanhos e fins diferentes e todos se concentram no objetivo de fornecer subsídios para o diagnóstico do comportamento da barragem que está comprometida pela ação do tempo e da ampla quantidade de água que a mesma tende a suportar. De acordo com Brasil (2002), toda barragem deve ser instrumentada adequadamente ao seu porte e aos riscos associados.

O instrumento conhecido como pêndulo simples consiste de “fio de prumo” gigante e é o maior instrumento da barragem e mede os deslocamentos horizontais dos blocos em vários pontos (cotas). O pêndulo direto afere o movimento horizontal da barragem com relação à fundação da estrutura e o invertido mede o deslocamento da fundação. De acordo com Silveira (2013), o pêndulo direto é um dos instrumentos de maior precisão com alta robustez e precisão, que deve ser instalado em barragens de concreto com altura mínima de 30 metros. Este será um dos instrumentos utilizados na modelagem das séries temporais univariadas e multivariadas, seja pela confiabilidade de suas aferições, seja pela quantidade de informações disponíveis.

Dada a magnitude da obra, a instrumentação de toda a estrutura é algo imprescindível. Há um total de 2384 instrumentos instalados no concreto e na

fundação, os quais permitem o acompanhamento do desempenho destas estruturas. A Figura 9 ilustra uma visão interior de um bloco e os instrumentos nele instalados.

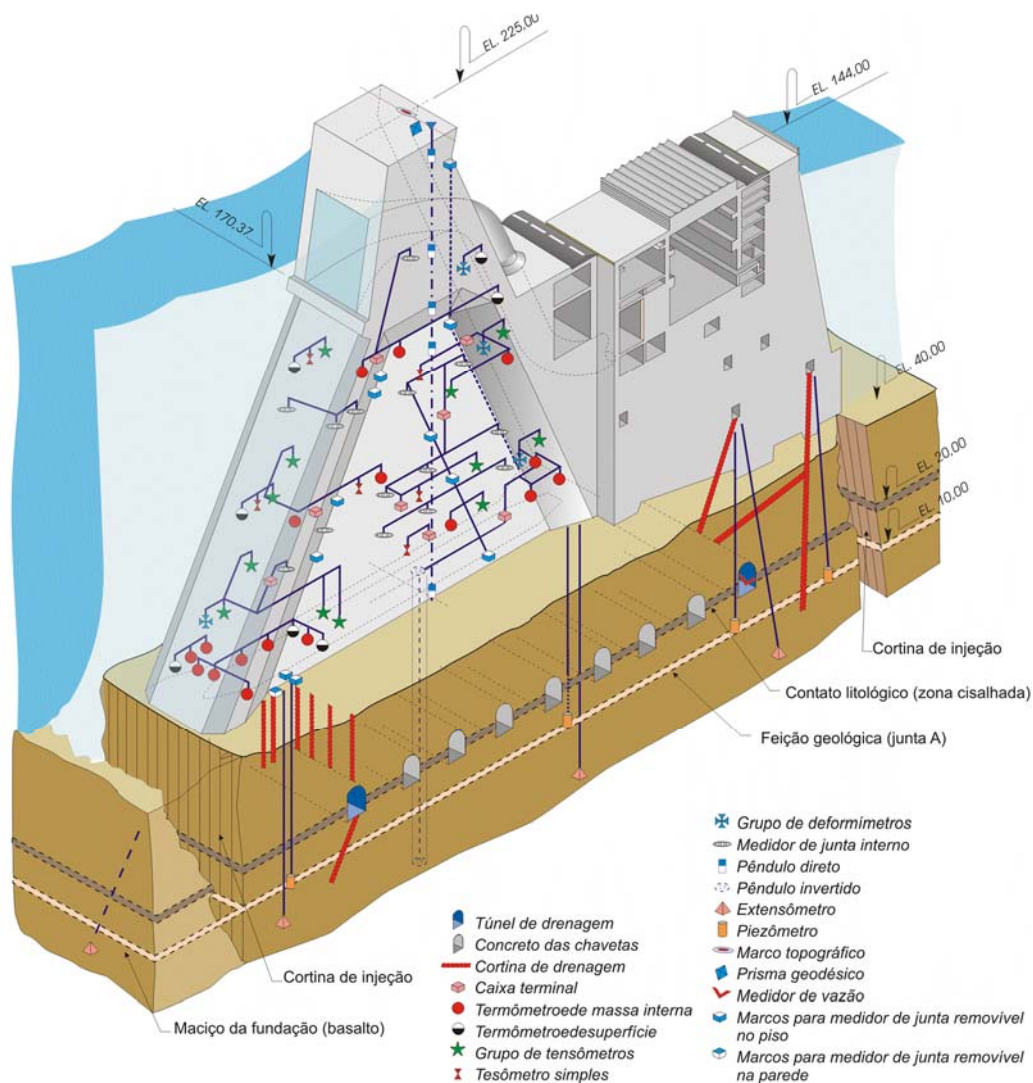


Figura 9 – Instrumentação

Fonte: (Itaipu (2015b))

Complementarmente, técnicos e engenheiros da Itaipu Binacional fazem o uso de técnicas topográficas e geodésicas, em que é possível mensurar diversos ângulos e distâncias por meio de equipamentos de ponta como a estação total TM30 (Figura 10). Este equipamento é instalado, a cada seis meses, em nove pilares e apontada para alvos, denominados prismas de

precisão (Figura 11). Estes alvos estão localizados no vertedouro, barragem lateral, barragem principal e barragem de enrocamento. Assim como o pêndulo, sua função é determinar a posição e, conseqüentemente, o deslocamento de cada um dos pontos inseridos da crista da barragem (alvos geodésicos). A cada período de coleta das informações em intervalos semestrais, denominado campanha, antes das aferições a estação total TM30 passa por rigorosos testes de calibração a fim de verificar as condições de uso.



Figura 10 - Estação total TM30



Figura 11 – Prismas de precisão localizados na barragem principal

A rede geodésica da Usina Hidrelétrica de Itaipu pode ser compreendida a partir do esquema da Figura 12.

Dessa forma, o estudo do deslocamento horizontal da barragem principal é de extrema importância para os engenheiros e técnicos da Itaipu, já que se trata de uma obra de grande porte do ponto de vista da engenharia. Os resultados das previsões das séries temporais de pêndulos, a partir da metodologia proposta, agregarão informações importantes aos valores limites estabelecidos no projeto de construção da usina. Com isso, será possível a adoção de melhorias estratégicas tanto em relação ao monitoramento do desempenho estrutural da usina, como em intervenções apropriadas, tendo como consequência, a redução de custo financeiro e de tempo.

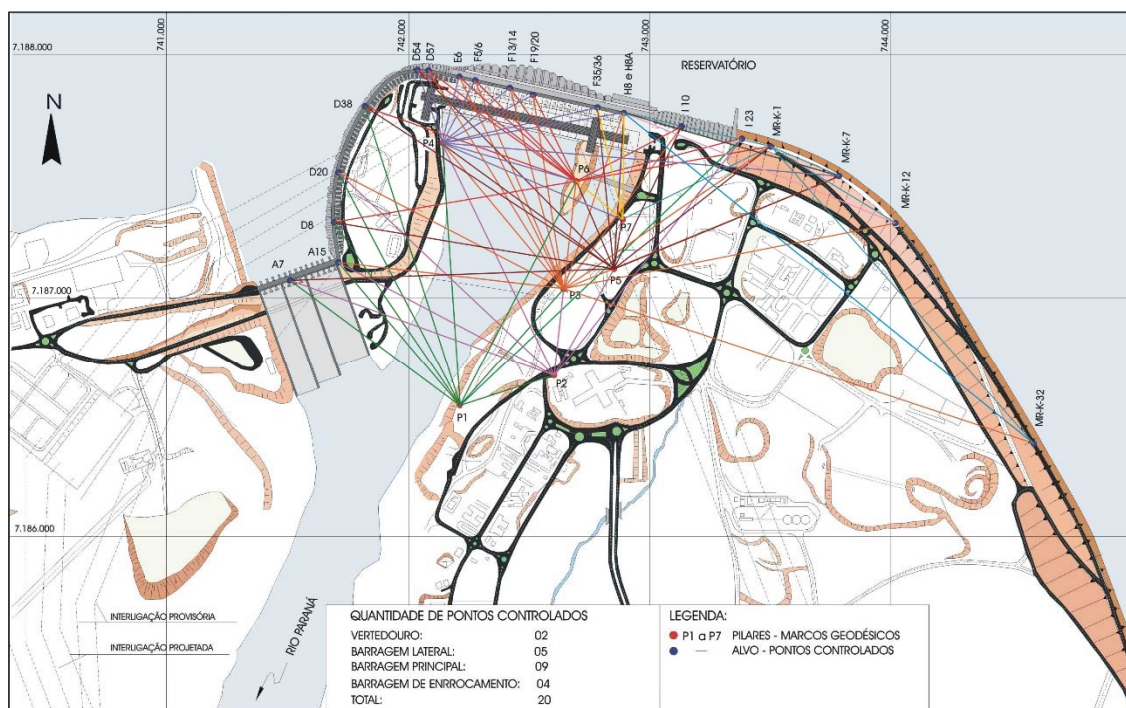


Figura 12 – Rede Geodésica de Itaipu
Fonte: (Itaipu (2015c))

As séries temporais utilizadas foram aquelas disponibilizadas pela Itaipu Binacional, as quais aferem os deslocamentos horizontais dos blocos F05, F13 e F19, localizados respectivamente, nas cotas 145,42m; 145,39m e 145,45m, na barragem principal da usina hidrelétrica. Estas séries temporais foram geradas a partir do coordenômetro automatizado. A seguir, é mostrado na Tabela 1, o conjunto das séries temporais utilizadas neste trabalho de tese, bem como suas cardinalidades e a natureza de cada uma delas.

Tabela 1 - Séries de deslocamentos dos blocos-chave da barragem principal da Usina Hidrelétrica de Itaipu com base nos pêndulos automatizados

Série	Cardinalidade	Aferição	Unidade de medida
F05x	2554	Deslocamento do bloco F05 sentido montante-jusante	Milímetros
F05y	2046	Deslocamento do bloco F05 sentido perpendicular a montante-jusante	Milímetros
F13x	2023	Deslocamento do bloco F13 sentido montante-jusante	Milímetros
F13y	2023	Deslocamento do bloco F13 sentido perpendicular a	Milímetros

montante-jusante			
F19x	2023	Deslocamento do bloco F19 sentido montante-jusante	Milímetros

Conforme metodologia WARIMAX-GARCH Neural, descrita na seção 3.2, as previsões foram obtidas a partir da decomposição *Wavelet*, modelagem das componentes *Wavelet* via ARIMA-GARCH (geração das séries artificiais ou sintéticas) e modelagem da série original via ARIMAX-GARCH com a entrada de séries sintéticas, dando origem ao modelo WARIMAX-GARCH.

Em princípio, as séries foram decompostas ortogonalmente via *Wavelet* com o uso do software MATLAB 2013. Cada uma das componentes *Wavelet* (aproximação e detalhe) foi obtida com o uso das bases ortonormais como Haar, Daubechies, Symlets, dentre outras, conforme o melhor ajuste na modelagem ARIMA-GARCH. A modelagem tanto das séries decompostas como da série original foi feita a partir do software EViews8.

3.2 Método Proposto

O método proposto, designado por WARIMAX-GARCH Neural, requer que as variáveis exógenas estejam disponíveis para serem utilizadas. No caso em que as variáveis exógenas não estejam disponíveis, um modelo ARIMA-GARCH pode ser utilizado, porém com ganho preditivo inferior. O método WARIMAX-GARCH Neural é definido a seguir, e se destina a resolver este problema (pelo menos em parte). É formado pelos passos detalhados mencionados abaixo.

Passo 1: É efetuada a Decomposição *Wavelet* (DW) de nível r (descrito na seção 2.1) da série temporal subjacente y_t ($t = 1, \dots, T$), gerando $r + 1$ componentes *Wavelet*. Isto é, uma componente *Wavelet* de aproximação de nível m_0 , denotada por $\tilde{y}_{A_{m_0},t}$ ($t = 1, \dots, T$), e r componentes *Wavelet* de detalhe de nível m_0 até $m_{0+(r-1)}$, denotada por $\tilde{y}_{D_{m,t}}$ ($t = 1, \dots, T$) para $m = m_0, \dots, m_{0+(r-1)}$.

Passo 2: Cada componente *Wavelet* obtida no passo 1 é modelada individualmente por meio de modelos ARIMA-GARCH e de RNA's, que são utilizados para produzirem as previsão fora da amostra.

Passo 3: Cada uma das componentes *Wavelet*, obtida no passo 1, é completada com previsões fora da amostra (com horizonte de previsão h) por ambos os métodos de previsão, do passo 2. A seguir, são combinadas linearmente e passam a ser utilizadas como variáveis exógenas *Wavelet* (VEW) no passo 4. Algebricamente, estas variáveis são representadas por:

- I. $\tilde{y}_{A_{m_0},t,C}$ ou $x_{1,t}$ ($t = 1, \dots, T, T+1, \dots, T+h$) que consistem de componentes *Wavelet* de aproximação de nível m_0 da série y_t ($t = 1, \dots, T$) e é tal que: $\tilde{y}_{A_{m_0},t,C} = \tilde{y}_{A_{m_0},t}$, se ($t = 1, \dots, T$); e $\tilde{y}_{A_{m_0},t,C} = \hat{y}_{A_{m_0},t}$ se ($t = T+1, \dots, T+h$), onde $\hat{y}_{A_{m_0},t}$ representa a previsão fora da amostra no instante t gerado pela combinação linear das previsões obtidas por ARIMA-GARCH e RNA no passo 2.
- II. $\tilde{y}_{D_m,t,C}$ ou $x_{i,t}$ ($t = 1, \dots, T, T+1, \dots, T+h$), onde $i = 2, \dots, (r+1)$, consistem de componentes *Wavelet* de detalhe de nível m , onde $m_0 \leq m \leq m_0 + (r-1)$, de y_t ($t = 1, \dots, T$) e é tal que: $\tilde{y}_{D_m,t,C} = \tilde{y}_{D_m,t}$, se $t = 1, \dots, T$; e $\tilde{y}_{D_m,t,C} = \hat{y}_{D_m,t}$, para $t = T+1, \dots, T+h$, onde $\hat{y}_{D_m,t}$ representa a previsão fora da amostra produzida num instante t pela combinação linear das previsões obtidas por ARIMA-GARCH e RNA no passo 2.

Passo 4: as $r+1$ componentes *Wavelet* completadas no passo 3 são tratadas como $r+1$ variáveis exógenas *Wavelet* utilizadas para o modelo WARIMAX-GARCH Neural $(p,d,q) \times (P,D,Q) \times (R,V)$ para gerar previsões dentro e fora da amostra para a média e a variância condicional da série y_t ($t = 1, \dots, T$), bem como os seus intervalos de previsão, no pressuposto dos resíduos fora da amostra serem gaussianos.

O modelo WARIMAX-GARCH Neural $(p,d,q) \times (P,D,Q) \times (R,V)$ é, formalmente, definido por (32) e (33), conforme segue:

$$\phi(B) \Phi(B^S) \nabla^d \nabla_S^D y_t = \varphi_{A_{m_0}} \nabla^k g(\tilde{y}_{A_{m_0},t,C}) + \sum_{m=m_0}^{m_0+(r-1)} \nabla^{k_m} \tilde{y}_{D_m,t,C} \varphi_{D_m,C} + \theta(B) \Theta(B^S) e_t + \varphi \sigma_t^2 \quad (32)$$

onde as componentes $\phi(B) := (1 - \phi_1 B - \dots - \phi_p B^p)$, $\theta(B) := (1 - \theta_1 B - \dots - \theta_q B^q)$, $\nabla^d := (1 - B)^d$, $\nabla_S^D := (1 - B^S)^D$, $\theta(B^S) := (1 - \theta_1 B^S - \dots - \theta_Q B^{SQ})$ e $\Phi(B^S) := (1 - \phi_1 B^S - \dots - \phi_Q B^{SQ})$ são os polinômios associados com o convencional modelo SARIMA $(p, d, q) \times (P, D, Q)$; φ é o parâmetro linearmente associado com a variância σ_t^2 ; $\varphi_{A_{m_0}}$ e $\varphi_{D_m,C}$ são os parâmetros linearmente associados com $\tilde{y}_{A_{m_0},t,C}$ e com $\tilde{y}_{D_m,t,C}$, respectivamente; $g(\cdot)$ é uma transformação que pode ser tanto a identidade quanto o mapeamento por logaritmo (que é muito útil, em diversos casos, para se obter estacionariedade na média, conforme Hamilton (1994)) bem como para melhorar o desempenho de previsão; ∇^k , onde $k \geq 0$ e $k \in \mathbb{Z}$, é o operador diferença associado com a variável exógena $\tilde{y}_{A_{m_0},t,C}$ que, conforme Hamilton (1994), pode ser usada para gerar uma nova série temporal com média constante (estacionária). É importante ressaltar que ∇^d e ∇^k podem assumir diferentes ordens quando se está procurando o modelo ideal.

No que diz respeito aos parâmetros do modelo WARIMAX, são exigidos os mesmos pressupostos dos parâmetros do ARIMAX convencional e são estimados utilizando o método da máxima verossimilhança. A fim de obter a variância do modelo para y_t ($t = 1, \dots, T$), o modelo WARIMAX-GARCH Neural usa as componentes do modelo GARCH (R, V) , que são dadas por

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^R \alpha_i e_{t-i}^2 + \sum_{j=1}^V \beta_j \sigma_{t-j}^2, \quad (33)$$

Onde $e_t := \sigma_t z_t$, com z_t sendo a realização de uma variável randômica Z_t , padronizada, independente e identicamente distribuída, σ_t é o desvio padrão condicional.

Passo 5: Uma vez que são geradas as previsões dentro e fora da amostra de \hat{y}_t ($t = 1, \dots, T, T+1, \dots, T+h$) para a média da série temporal y_t ($t =$

$1, \dots, T$), e a previsão para sua variância (volatilidade) $\hat{\sigma}_t^2$ ($t = 1, \dots, T, T + 1, \dots, T + h$) é determinada, o intervalo de previsão $[\hat{y}_{INF,t}; \hat{y}_{SUP,t}]$ ($t = 1, \dots, T, T + 1, \dots, T + h$), com nível de confiança $1 - \alpha$, sob a suposição de resíduos gaussianos, pode ser trivialmente determinado. O limite inferior $\hat{y}_{INF,t}$ e o limite superior $\hat{y}_{SUP,t}$ dos intervalos de previsão no instante t são, respectivamente, definidos por: $\hat{y}_{INF,t} := \hat{y}_t - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ e $\hat{y}_{SUP,t} := \hat{y}_t + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$, onde $z_{\frac{\alpha}{2}}$ é a realização da variável randômica i.i.d. Z_t associada a α .

Nota-se que embora os métodos *Wavelet* sejam aplicáveis em séries temporais não estacionárias e/ou não lineares (Mallat, 2008), todas as variáveis exógenas *Wavelet* do modelo WARIMAX-GARCH Neural devem ser estacionárias, pois devem satisfazer tal requisito para o modelo ARIMAX. No entanto, variáveis exógenas *Wavelet* não estacionárias, sendo uma componente *Wavelet* de aproximação $\tilde{y}_{A_{m_0},t,c}$ ($t = 1, \dots, T$), bem como suas versões defasadas, podem se tornar estacionárias via processo de diferenciação com o uso do operador de atraso ∇^k e/ou por transformação por meio do mapeamento $g(\cdot)$ associado com a componente, assim como é feito na modelagem ARIMA quando a série original não é estacionária.

Observa-se também que, as variáveis exógenas *Wavelet* que consistem de componentes *Wavelet* de detalhe, são sempre estacionárias na média tal que o operador diferença ∇^{k_m} em (32) é aplicado apenas para ajudar a alcançar melhores performances de previsão e/ou obter um modelo plausível. De fato, conforme Mallat (2008), a função *Wavelet* de nível m , $\omega_{m,n}(\cdot)$, consiste de uma curva de curta duração cujos valores da imagem têm média zero. Além disso, uma componente *Wavelet* de detalhe de nível m , $\tilde{y}_{D_{m,t}}$ ($t = 1, \dots, T$), é definida matematicamente pela combinação linear de funções *Wavelet* de nível m tal que $\tilde{y}_{D_{m,t}} := \sum_{n=0}^{2^{(M-m_0)}-1} d_{m,n} \omega_{m,n}(t)$ para cada instante t , onde $t = 1, \dots, T$. Similarmente para $\omega_{m,n}(\cdot)$, a componente *Wavelet* de detalhe $\tilde{y}_{D_{m,t}}$ assume valores em torno de zero, do ponto de vista estatístico, isso significa que $\tilde{y}_{D_{m,t}}$ exibe estacionariedade na média condicional. Agora, baseado em Mallat

(2008), desde que m seja um parâmetro fixo associado com a frequência espectral de $\tilde{y}_{D_m,t}$, segue que as variâncias condicionais (volatilidades) têm flutuações estocásticas estacionárias.

O fluxograma da Figura 13 apresenta a estrutura geral deste trabalho, com os passos a serem desenvolvidos, conforme descritos no início desta seção.

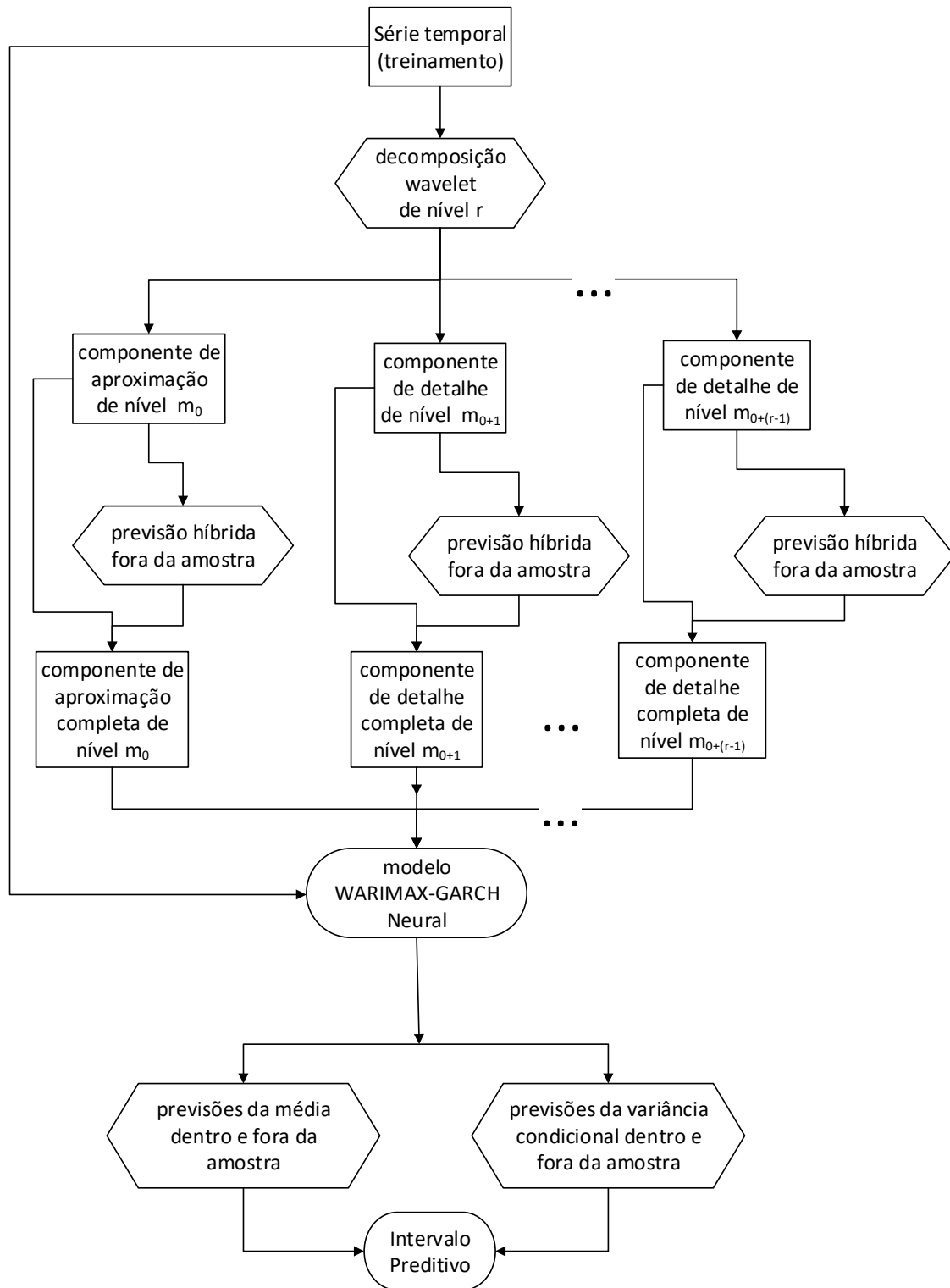


Figura 13 - Fluxograma com etapas do método proposto

O fluxograma da Figura 14 detalha a etapa da “previsão híbrida fora da amostra” relacionada ao fluxograma da Figura 13.

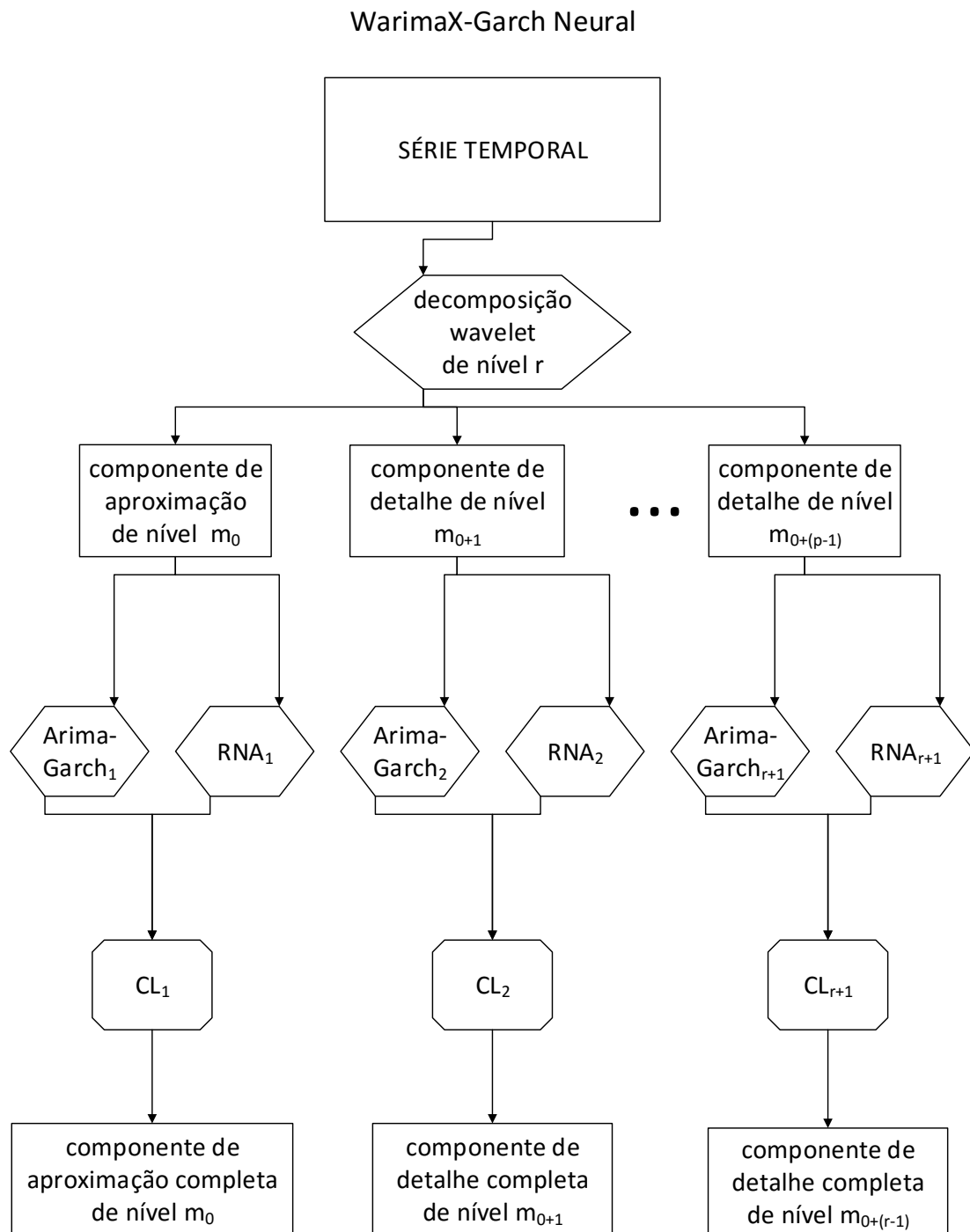


Figura 14 – Geração das variáveis exógenas híbridas

Vale salientar que o modelo WARIMAX-GARCH foi designado dessa forma quando não houve uso de RNA para a modelagem das componentes *Wavelet* (Experimento Computacional da seção 4.1).

4 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os experimentos computacionais e as considerações relativas aos resultados obtidos com o uso da metodologia proposta.

4.1 Experimento Computacional: Método WARIMAX-GARCH

Neste experimento foi utilizada a série temporal F05x, descrita na Tabela 1 (página 60), que é composta de 2554 valores aferidos diariamente às 10h00 da manhã, entre os dias 28/10/2005 e 24/10/2012, de forma automatizada. Foram utilizadas 2506 observações (dias) para treino e 48 observações (dias) foram utilizadas para teste, ou seja, horizonte de previsão $h = 48$ com 48 passos à frente. O gráfico da série temporal é apresentado na Figura 15.

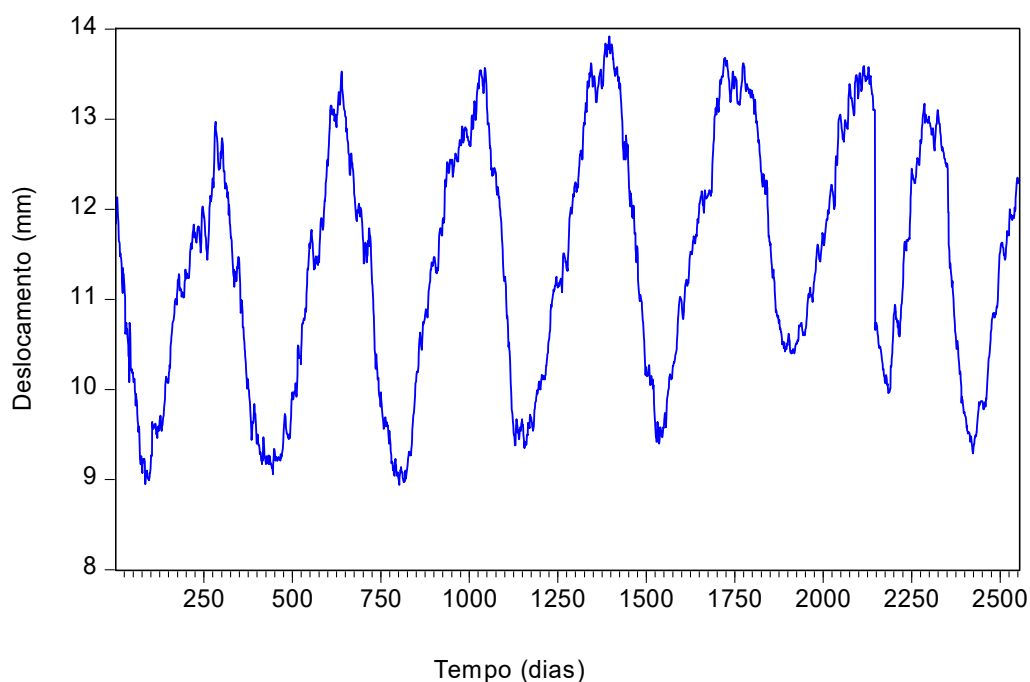


Figura 15 – Série temporal do Deslocamento do bloco F05x

Nota-se na Figura que a série temporal apresenta oscilações de alta e baixa frequência nos dados do deslocamento. Embora um modelo ARIMA sazonal seja apto a capturar as variações de baixa frequência da série temporal subjacente, em alguns casos, pode não ser capaz de modelar, de forma eficiente, as variações de alta frequência (representadas pelas componentes *Wavelet* de detalhes). Um modelo WARIMAX-GARCH, porém, pode modelar ambas (isto é, as oscilações de alta e baixa frequência) no domínio do tempo da série temporal subjacente, através dos modelos ARIMA-GARCH integrados com a abordagem de decomposição *Wavelet*.

O teste da raiz unitária Dickey-Fuller Aumentado (DFA) foi realizado e não rejeitou ao nível de significância de 1%, a hipótese de não estacionariedade dos dados de deslocamento. No entanto, o teste DFA rejeitou fortemente a não estacionariedade da série temporal mapeada com a função logaritmo, ao nível de 1%, bem como a segunda diferença desta série. A segunda diferença, aqui, não foi utilizada para obter estacionariedade, mas para obter maior aderência do modelo à série. As funções de autocorrelação simples e parciais, em relação a este último, bem como sua versão ao quadrado, apresentam valores significativos ao nível de 1%, o que sugere a escolha de um modelo ARIMA-GARCH. Portanto, uma vez que não houve variáveis exógenas convencionais disponíveis para serem utilizadas pelo modelo ARIMAX-GARCH nesta aplicação, foi apresentado um modelo ARIMA-GARCH que foi considerado plausível para o processo de modelagem.

A investigação conduzida nesta seção gerou previsões pontuais e intervalares utilizando a metodologia 1-passo dentro da amostra e multipassos fora da amostra, tanto para o modelo ARIMA-GARCH (usado como referência) quanto para o método WARIMAX-GARCH. Diagnósticos residuais para a validação estatística foram realizados usando o teste DFA, testes de Ljung-Box e Durbin-Watson para autocorrelação de primeira ordem, bem como os gráficos das funções de autocorreção e autocorrelação parciais, teste estatístico BDS para detectar não linearidade nas autocorrelações da série e o teste Arch-LM

para a variância incondicionalmente constante (HAMILTON, 1994). O Software EViews 8 foi utilizado para realizar os ensaios acima.

4.1.1 Modelo ARIMA-GARCH

Um modelo ARIMA (3,2,4) integrado com o modelo GARCH (2,1) (isto é, modelo ARIMA-GARCH (3,2,4) x (2,1)), com Distribuição do Erro Generalizada, (GED) (LIU, 2006) foi identificado para produzir o melhor ajuste para a amostra de treinamento y'_t ($t = 1, \dots, 2506$), duas vezes diferenciada, com mapeamento com a função logaritmo. Os parâmetros deste modelo foram estimados por Máxima Verossimilhança e foram estatisticamente significativos ao nível de 1% (ver Apêndice I B).

Entre todos os modelos ARIMA-GARCH plausíveis obtidos a partir da amostra de treinamento, o modelo ARIMA-GARCH (3,2,4) x (2,1) produziu as previsões mais precisas para a média da série histórica subjacente dentro da amostra de treinamento, em termos das estatísticas MAPE e MAE. Os resultados dos testes de Ljung-Box (Q-Stat) aplicados aos resíduos ordinários do modelo ARIMA-GARCH (3,2,4) x (2,1), conforme Figura 16 e Figura 17, sugerem que não há estrutura de autodependência linear significativa (ao nível de 1%) em ambos os resíduos, ordinários e quadráticos, até o $lag = 36$ (correspondendo a 36 dias). Além disso, o cálculo da estatística de Durbin-Watson, com valor 2,092367, sugere que não há autocorrelação de $lag = 1$ significativa nos resíduos ordinários (Maiores detalhes ver Apêndice I B e Hamilton (1994), para maiores informações sobre as técnicas estatísticas).

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	-0.021	-0.021	1.0781		
2	0.013	0.013	1.5119		
3	0.001	0.002	1.5164		
4	0.025	0.025	3.1276		
5	0.027	0.028	5.0089	0.025	
6	0.033	0.033	7.7102	0.021	
7	0.021	0.022	8.8475	0.031	
8	0.007	0.006	8.9698	0.062	
9	0.026	0.025	10.707	0.058	
10	0.020	0.019	11.760	0.068	
11	0.034	0.031	14.587	0.042	
12	0.021	0.020	15.748	0.046	
13	0.013	0.011	16.201	0.063	
14	0.025	0.022	17.754	0.059	
15	0.012	0.009	18.124	0.079	
16	0.012	0.007	18.483	0.102	
17	0.029	0.025	20.638	0.080	
18	0.037	0.033	24.051	0.045	
19	0.032	0.029	26.669	0.032	
20	0.013	0.010	27.129	0.040	
21	0.026	0.021	28.873	0.036	
22	0.026	0.021	30.529	0.033	
23	0.010	0.003	30.781	0.043	
24	0.029	0.021	32.842	0.035	
25	0.005	-0.002	32.895	0.047	
26	0.018	0.011	33.746	0.052	
27	0.031	0.024	36.139	0.040	
28	0.028	0.021	38.154	0.033	
29	0.020	0.014	39.200	0.035	
30	0.027	0.020	41.042	0.031	
31	0.033	0.026	43.726	0.022	
32	0.015	0.008	44.288	0.026	
33	0.036	0.027	47.617	0.016	
34	0.010	0.003	47.871	0.020	
35	0.001	-0.009	47.872	0.027	
36	0.007	-0.003	48.004	0.034	

Figura 16 – FAC e FACP dos resíduos ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	-0.001	-0.001	0.0019	0.965	
2	-0.001	-0.001	0.0044	0.998	
3	-0.001	-0.001	0.0074	1.000	
4	-0.001	-0.001	0.0093	1.000	
5	-0.001	-0.001	0.0118	1.000	
6	-0.001	-0.001	0.0125	1.000	
7	-0.001	-0.001	0.0143	1.000	
8	-0.001	-0.001	0.0172	1.000	
9	-0.001	-0.001	0.0199	1.000	
10	-0.001	-0.001	0.0230	1.000	
11	-0.001	-0.001	0.0252	1.000	
12	-0.000	-0.000	0.0255	1.000	
13	-0.001	-0.001	0.0283	1.000	
14	-0.001	-0.001	0.0314	1.000	
15	-0.000	-0.000	0.0317	1.000	
16	-0.001	-0.001	0.0345	1.000	
17	-0.001	-0.001	0.0357	1.000	
18	-0.000	-0.000	0.0361	1.000	
19	-0.001	-0.001	0.0381	1.000	
20	-0.001	-0.001	0.0402	1.000	
21	-0.001	-0.001	0.0434	1.000	
22	-0.001	-0.001	0.0463	1.000	
23	-0.001	-0.001	0.0491	1.000	
24	-0.001	-0.001	0.0504	1.000	
25	-0.001	-0.001	0.0527	1.000	
26	-0.001	-0.001	0.0552	1.000	
27	-0.001	-0.001	0.0577	1.000	
28	-0.001	-0.001	0.0602	1.000	
29	-0.001	-0.001	0.0627	1.000	
30	-0.001	-0.001	0.0646	1.000	
31	-0.001	-0.001	0.0657	1.000	
32	-0.001	-0.001	0.0685	1.000	
33	-0.001	-0.001	0.0709	1.000	
34	-0.001	-0.001	0.0737	1.000	
35	-0.001	-0.001	0.0762	1.000	
36	-0.000	-0.000	0.0767	1.000	

Figura 17 – FAC e FACP dos resíduos quadráticos padronizados

<i>dimensions</i>	<i>BDS statistics</i>	<i>p-values</i>
2	-3.22E-07	0.9840
3	-9.65E-07	0.9785
4	-1.93E-06	0.9743
5	-3.22E-06	0.9707
6	-4.83E-06	0.9675

Tabela 2 – Teste BDS dentro da amostra de teste para os resíduos ordinários do modelo ARIMA-GARCH

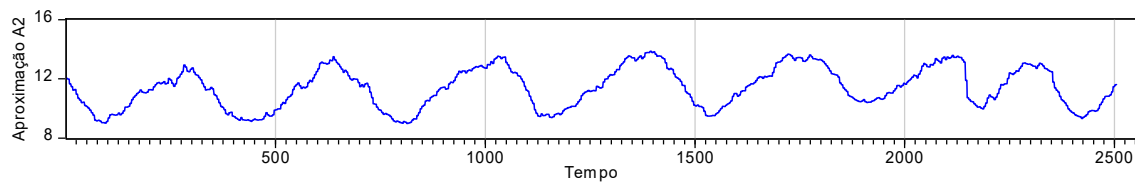
A Tabela 2 mostra as estatísticas e os p-valores correspondentes para dimensões de 2 a 5 do teste BDS (que consiste de um teste estatístico para verificar a existência de autodependência linear e não linear em um conjunto de dados (HAMILTON, 1994)) aplicado na análise residual (dentro da amostra de treinamento) do modelo ARIMA-GARCH (3,2,4) x (2,1). De acordo com estes valores de p, é possível concluir que não há evidência, ao nível de 1% de significância e em todas as cinco dimensões, de autodependência linear e não linear nos erros de previsão dentro da amostra. Em particular, as estruturas de autodependência linear e não linear anteriormente presentes nos resíduos de previsão foram adequadamente mapeadas pelos modelos ARIMA-GARCH

especificados acima. Portanto, os resíduos de previsão dentro da amostra podem ser considerados como um ruído branco com média zero.

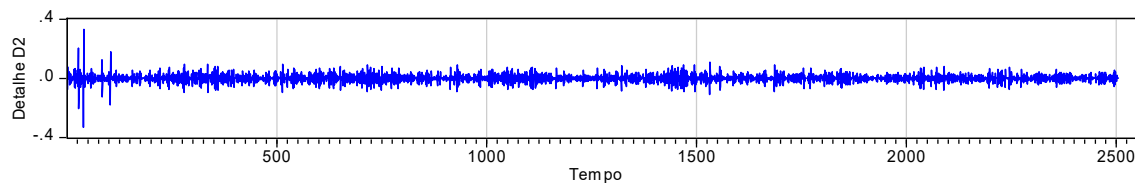
Vale ressaltar que a escolha do modelo ARIMA-GARCH, entre todos os modelos plausíveis identificados, foi determinada pela comparação do desempenho de previsão com o uso das estatísticas APE, MAPE, MAE e R^2 , em cada modelo candidato. O modelo selecionado foi aquele com menor APE, MAPE e MAE e com o maior R^2 .

4.1.2 Modelo WARIMAX-GARCH: Série F05x

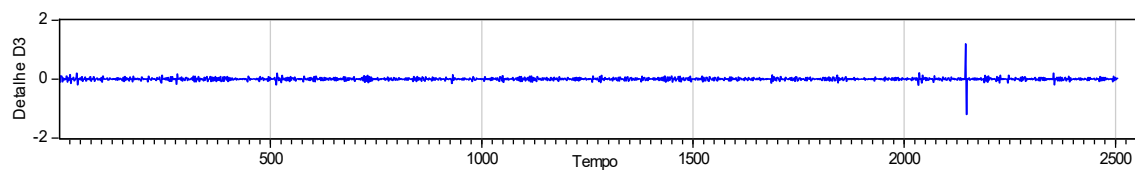
Num primeiro instante, o método WARIMAX-GARCH foi implementado no software MATLAB (versão 2013a). A decomposição *Wavelet* de nível $r = 2$ foi obtida a partir da amostra de treino da série de deslocamento, no sentido montante-jusante, do bloco F05 da barragem principal da Usina Hidrelétrica de Itaipu. Os gráficos das componentes *Wavelet*, com base ortonormal *db40* (DAUBECHIES, 1988), podem ser vistos na Figura 18.



(a) Componente *Wavelet* de aproximação de nível 2, $\tilde{y}_{A_2,t}$ ($t = 1, \dots, 2506$)



(b) Componente *Wavelet* de detalhe de nível 2, $\tilde{y}_{D_2,t}$ ($t = 1, \dots, 2506$)



(c) Componente *Wavelet* de detalhe de nível 3, $\tilde{y}_{D_3,t}$ ($t = 1, \dots, 2506$)

Figura 18 – Componentes *Wavelet* da amostra de treino da série F05x

De acordo com a seção 2.1, para esta aplicação, $y_t = \tilde{y}_{A_2,t} + \tilde{y}_{D_2,t} + \tilde{y}_{D_3,t} + \varepsilon$, ($t = 1, \dots, 2506$) onde ε é, estatisticamente, considerado ruído branco. Numa segunda etapa do método WARIMAX-GARCH, cada uma das componentes *Wavelet* $\tilde{y}_{A_2,t}$, $\tilde{y}_{D_2,t}$ e $\tilde{y}_{D_3,t}$ ($t = 1, \dots, 2506$) foi modelada, individualmente, por três modelos ARIMA-GARCH apropriados para cada série e, com estes modelos, foram geradas as previsões fora da amostra para suas respectivas médias, denotadas por $\hat{y}_{A_2,t}$, $\hat{y}_{D_2,t}$ e $\hat{y}_{D_3,t}$ ($t = 2507, \dots, 2554$). O horizonte de previsão escolhido foi $h=48$, com 48 passos à frente, ou seja, foram previstos 48 dias de deslocamento do bloco F05 (sentido montante-jusante), dado pela série F05x. Todos os parâmetros dos três modelos estimados são estatisticamente significativos, a 1%, e o diagnóstico residual confirma sua plausibilidade (Ver Apêndice I B).

Numa terceira etapa, as três componentes *Wavelet* $\tilde{y}_{A_2,t,C}$, $\tilde{y}_{D_2,t,C}$ e $\tilde{y}_{D_3,t,C}$ (ou x_1, x_2 e x_3 , respectivamente) consistem de variáveis exógenas *Wavelet* geradas a partir da série de deslocamento F05x. Como mencionado anteriormente, elas são facilmente obtidas completando as componentes *Wavelet* com suas previsões fora da amostra, produzidas na segunda etapa.

Algebricamente, são representadas por:

- I. $x_{1,t} = \tilde{y}_{A_2,t,C} := \left((\tilde{y}_{A_2,t})_{t=1}^{2506}; (\hat{y}_{A_2,t})_{t=2507}^{2554} \right);$
- II. $x_{2,t} = \tilde{y}_{D_2,t,C} := \left((\tilde{y}_{D_2,t})_{t=1}^{2506}; (\hat{y}_{D_2,t})_{t=2507}^{2554} \right);$ e
- III. $x_{3,t} = \tilde{y}_{D_3,t,C} := \left((\tilde{y}_{D_3,t})_{t=1}^{2506}; (\hat{y}_{D_3,t})_{t=2507}^{2554} \right);$

Numa quarta etapa, foram usadas as variáveis exógenas *Wavelet* $x_{i,t}$ ($i = 1, 2, 3$) ($t = 1, \dots, 2554$) para o ajustamento do modelo WARIMAX-GARCH utilizado para gerar as 48 previsões da média e da variância condicional da série F05x (com 48 passos à frente). Algebricamente, a melhor formulação do modelo WARIMAX obtida foi:

$$(1 - \phi_1 B - \dots - \phi_5 B^5) y_t = (1 - \theta_1 B - \dots - \theta_5 B^5) e_t + \sum_{i=1}^2 \tilde{y}_{A_2,t-i,C} \varphi_{A_2,t-i} + \tilde{y}_{D_2,t-1,C} \varphi_{D_2,t-1} + \tilde{y}_{D_3,t-2,C} \varphi_{D_3,t-2} + \varphi \sigma_t^2$$

No Apêndice I B, o modelo WARIMAX acima, em conjunto com as principais estatísticas, pode ser visto com maiores detalhes. A volatilidade σ_t^2 proporciona um efeito não linear para representar o estado y_t . Também, o melhor modelo GARCH obtido foi GARCH (1,1), com distribuição do erro generalizada (GED), com a seguinte formulação algébrica:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

O método da Máxima Verossimilhança foi utilizado para obter as melhores estimativas dos parâmetros WARIMAX-GARCH. Os resultados apresentados no Apêndice I C mostram que todos os parâmetros do modelo WARIMAX-GARCH foram considerados estatisticamente significativos ao nível de 1%.

As Figura 19 e Figura 20 mostram, respectivamente, os gráficos da FAC e FACP (para lags de 1 até 36) dos resíduos ordinários padronizados e resíduos quadráticos padronizados associados ao modelo WARIMAX-GARCH. Nota-se que todos os valores da FAC e FACP pertencem ao intervalo de confiança com 99%, sugerindo que todos os valores são não significativos. O teste de Ljung-Box (Q-Stat) sugere, também, que não há autodependência linear significativa, ao nível de 1% (dentro da amostra), para os resíduos ordinários e quadráticos padronizados provenientes do modelo WARIMAX-GARCH. Também, o teste Ljung-Box para os resíduos quadráticos, na Figura 20, confirma que não há efeito ARCH nos resíduos da previsão. A estatística de Durbin-Watson, cujo valor obtido foi 2,015345 confirmou a ausência de autocorrelação de primeira ordem, dentro da amostra, para os resíduos padronizados (ver Apêndice I D).

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.008	-0.008	0.1531	
		2 -0.050	-0.050	6.3235	
		3 -0.016	-0.017	6.9561	
		4 -0.011	-0.014	7.2530	
		5 -0.001	-0.003	7.2569	
		6 0.006	0.005	7.3610	
		7 0.004	0.003	7.3970	
		8 -0.010	-0.010	7.6477	
		9 0.001	0.001	7.6512	
		10 0.022	0.022	8.9088	0.012
		11 0.007	0.007	9.0266	0.029
		12 -0.014	-0.012	9.5305	0.049
		13 -0.003	-0.002	9.5503	0.089
		14 -0.001	-0.001	9.5514	0.145
		15 -0.003	-0.003	9.5729	0.214
		16 0.017	0.016	10.304	0.244
		17 -0.002	-0.002	10.311	0.326
		18 0.034	0.036	13.227	0.211
		19 0.007	0.008	13.353	0.271
		20 0.032	0.035	15.929	0.195
		21 0.004	0.007	15.979	0.250
		22 0.005	0.010	16.036	0.311
		23 -0.004	-0.002	16.078	0.377
		24 0.003	0.004	16.097	0.446
		25 -0.002	-0.002	16.104	0.516
		26 0.028	0.028	18.115	0.448
		27 0.001	0.001	18.116	0.515
		28 0.017	0.019	18.842	0.532
		29 0.001	0.002	18.846	0.595
		30 0.046	0.049	24.202	0.337
		31 0.001	0.003	24.205	0.393
		32 0.029	0.035	26.329	0.337
		33 0.008	0.011	26.487	0.382
		34 -0.011	-0.006	26.773	0.421
		35 -0.007	-0.005	26.893	0.470
		36 -0.008	-0.011	27.042	0.516

Figura 19 - FAC e FACP dos resíduos ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.004	0.004	0.0343	0.853
		2 0.008	0.008	0.2130	0.899
		3 -0.002	-0.002	0.2233	0.974
		4 -0.000	-0.000	0.2235	0.994
		5 -0.003	-0.003	0.2396	0.999
		6 0.001	0.001	0.2433	1.000
		7 -0.003	-0.003	0.2594	1.000
		8 -0.000	-0.000	0.2599	1.000
		9 -0.003	-0.003	0.2763	1.000
		10 0.000	0.000	0.2768	1.000
		11 -0.002	-0.002	0.2922	1.000
		12 0.001	0.001	0.2931	1.000
		13 -0.003	-0.003	0.3097	1.000
		14 0.000	0.000	0.3103	1.000
		15 -0.003	-0.002	0.3261	1.000
		16 0.001	0.001	0.3311	1.000
		17 -0.003	-0.003	0.3476	1.000
		18 -0.000	-0.000	0.3481	1.000
		19 -0.002	-0.002	0.3637	1.000
		20 0.000	0.000	0.3640	1.000
		21 -0.003	-0.003	0.3810	1.000
		22 -0.000	-0.000	0.3811	1.000
		23 -0.003	-0.003	0.3986	1.000
		24 -0.001	-0.001	0.4002	1.000
		25 -0.003	-0.003	0.4171	1.000
		26 -0.000	-0.000	0.4171	1.000
		27 -0.003	-0.003	0.4341	1.000
		28 0.000	0.000	0.4342	1.000
		29 -0.003	-0.002	0.4503	1.000
		30 0.002	0.002	0.4642	1.000
		31 -0.003	-0.003	0.4808	1.000
		32 -0.001	-0.001	0.4816	1.000
		33 -0.003	-0.003	0.4983	1.000
		34 0.001	0.001	0.4991	1.000
		35 -0.003	-0.002	0.5155	1.000
		36 0.002	0.002	0.5232	1.000

Figura 20 - FAC e FACP dos resíduos quadráticos padronizados

Maiores detalhes da modelagem se encontram nos Apêndices desse trabalho (página 129).

Com base na Tabela 3, é possível comparar os resultados entre ARIMA-GARCH, *Wavelet*-RNA e WARIMAX-GARCH dentro e fora da amostra de

treino avaliando as estatísticas MAPE (*Mean Absolute Percent Error*) e MAE (*Mean Absolute Error*). Baseado nos dois tipos de erro, o modelo WARIMAX-GARCH teve melhor desempenho que o modelo ARIMA-GARCH (3, 2, 4) (2, 1), e houve redução de até 83% das estatísticas MAPE e MAE fora da amostra. Isso significa que o WARIMAX-GARCH modelou melhor a dinâmica da série e também teve melhor poder de gerar valores (previsões) fora da amostra.

MÉTODO	MAPE		MAE	
	<i>Dentro da amostra</i>	<i>Fora da amostra</i>	<i>Dentro da amostra</i>	<i>Fora da amostra</i>
ARIMA-GARCH	0,398%	5,324%	0,0448	0,6368
RNA	0,425%	2,247%	0,0474	0,2710
<i>Wavelet</i> -RNA	0,438%	1,840%	0,0495	0,2216
WARIMAX-GARCH	0,300%	0,899%	0,0338	0,1065

Tabela 3 – Valores dos erros MAPE e MAE, dentro e fora da amostra, para a série temporal F05X

A Figura 21 mostra o gráfico do Erro Percentual Absoluto (APE) calculado fora da amostra para cada passo à frente ($t = 2507, \dots, 2554$) para os modelos ARIMA-GARCH e WARIMAX-GARCH. Nota-se que o modelo WARIMAX-GARCH possui melhores valores APE que o modelo ARIMA-GARCH nos 48 valores fora da amostra, além de produzir previsões comparativamente melhores conforme o número de passos à frente vai aumentando. De fato, enquanto os erros de previsão do modelo ARIMA-GARCH mostram uma tendência positiva (crescente de pouco menos de 1% a quase 9%), o modelo WARIMAX-GARCH mostra flutuações ao longo da linha de 1% com o aumento do horizonte de previsão. O Apêndice I C mostra outros gráficos da evolução temporal fora da amostra e as respectivas previsões de ambas as abordagens para apoiar a conclusão de que, de fato, o modelo WARIMAX-GARCH tem melhor poder de generalização.

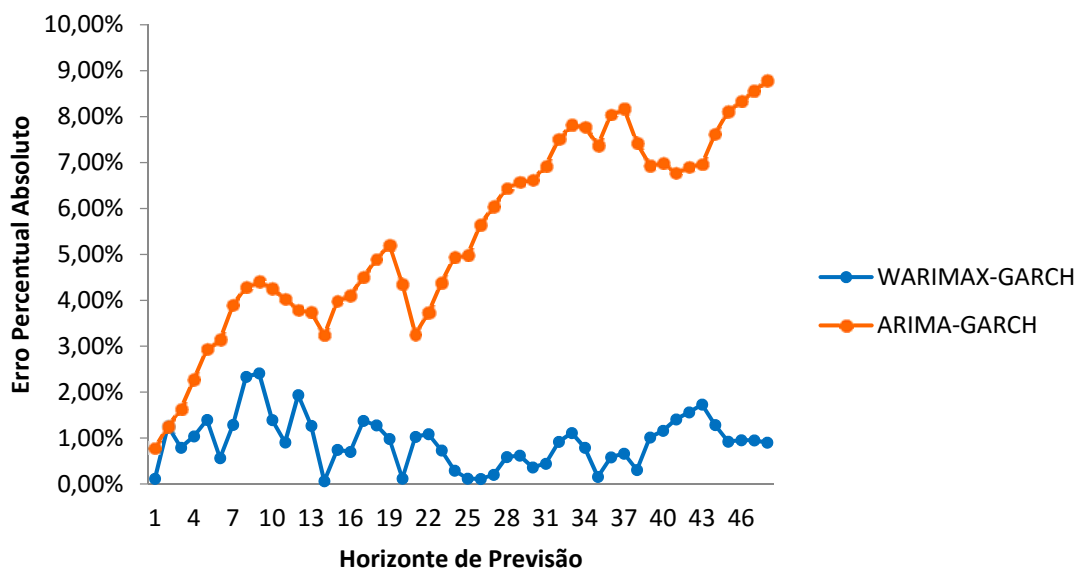


Figura 21 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH e ARIMA-GARCH

A Figura 22 exibe o gráfico dos erros percentuais absolutos calculado para as previsões em cada instante t fora da amostra ($t = 2507, \dots, 2554$) para cada passo à frente pelos métodos WARIMAX-GARCH e *Wavelet*-RNA, que obteve o segundo melhor resultado. A partir do passo 20, o método WARIMAX-GARCH, notavelmente, obteve melhor precisão das previsões que o método *Wavelet*-RNA, representando o melhor desempenho de previsão entre todos os métodos de referência.

Nota-se que, similarmente ao modelo ARIMA-GARCH, o método *Wavelet*-RNA tem perdido precisão na previsão para um horizonte de previsão maior, enquanto o método WARIMAX-GARCH manteve sua capacidade de previsão.

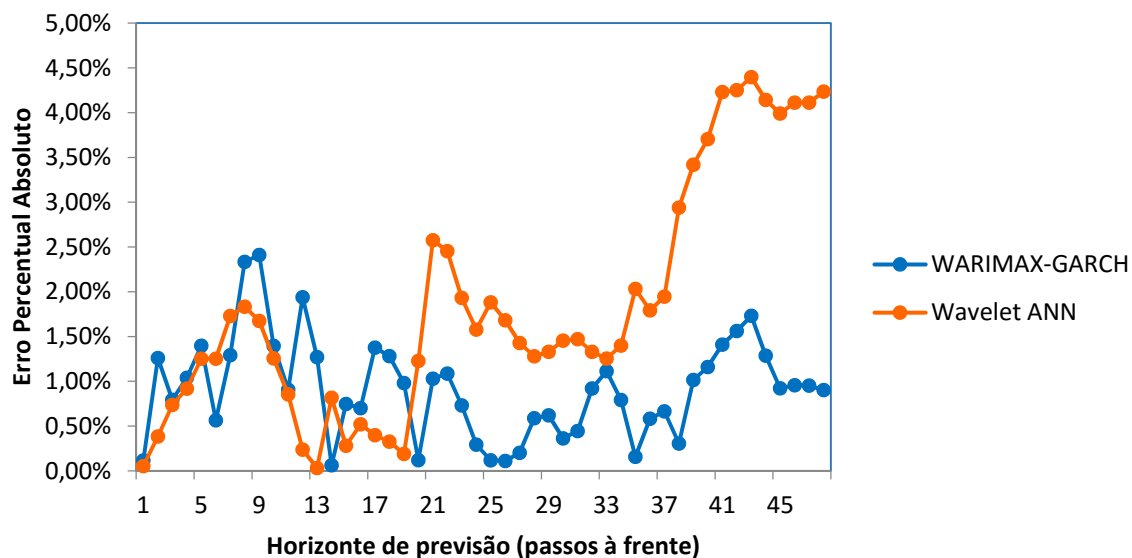


Figura 22 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH e RNA

A seguir são apresentados os gráficos dos deslocamentos reais e suas previsões, com intervalo de previsão de 99% para as previsões fora da amostra, obtidas por (a) Modelo ARIMA-GARCH (Figura 23) e (b) o Método WARIMAX-GARCH (Figura 24). Nota-se, pela Figura 23, que as previsões obtidas pelo modelo ARIMA-GARCH foram maiores do que os valores observados em cada momento, no entanto, todos os valores permaneceram dentro do intervalo de previsão de 99% para previsões fora da amostra.

Por outro lado, na Figura 24, as previsões obtidas pelo método WARIMAX-GARCH descreveram a trajetória do deslocamento mais próxima do deslocamento real e, também, todos os valores fizeram dentro do intervalo de 99% de confiança para a previsão, cujos limites foram muito menores que os do modelo ARIMA-GARCH. Isto é, a variância da densidade preditiva do método WARIMAX-GARCH foi menor. Este gráfico também mostra a dinâmica das previsões produzidas pelo método WARIMAX-GARCH por meio das oscilações das previsões em torno dos valores reais de deslocamento.

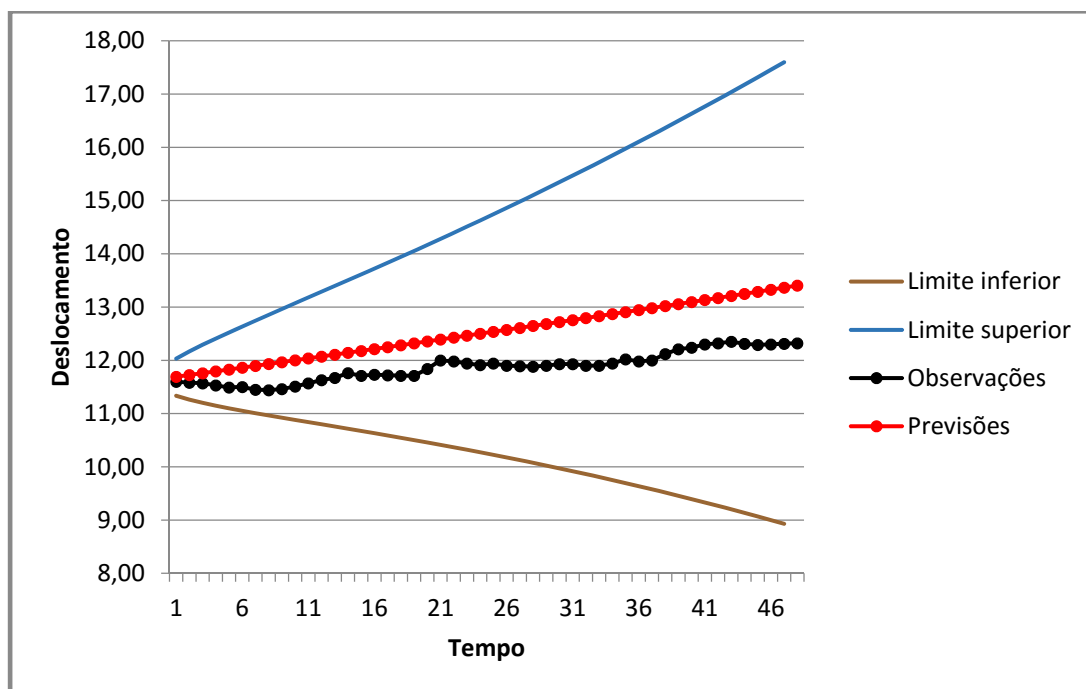


Figura 23 – Evolução temporal do intervalo de confiança para ARIMA-GARCH

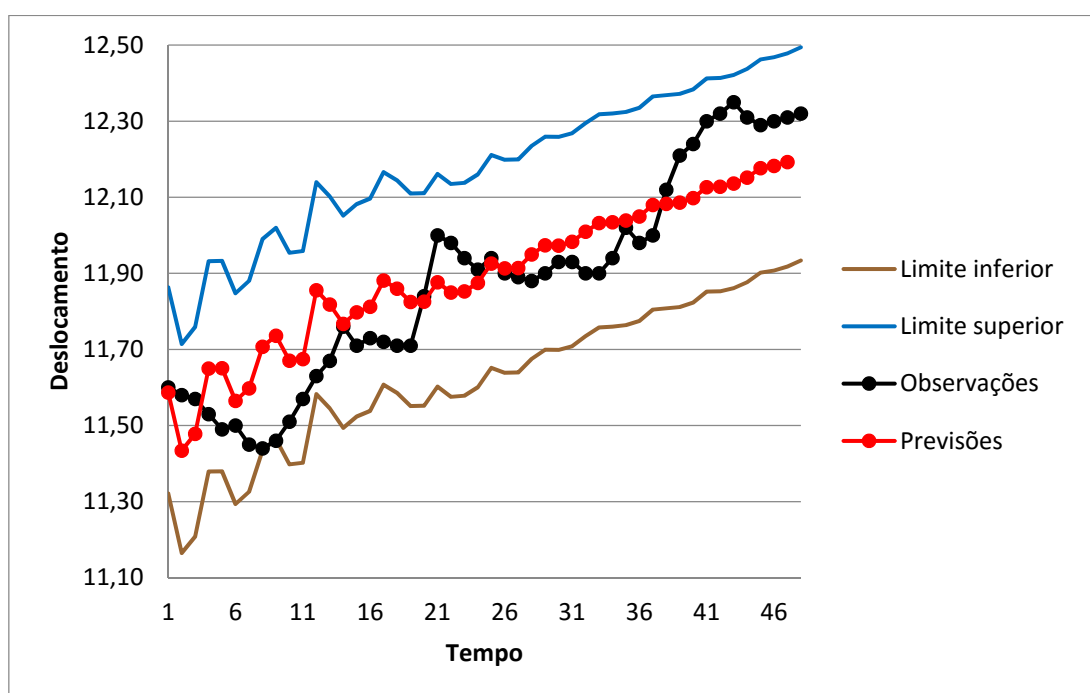


Figura 24 – Evolução Temporal do Intervalo de Confiança para WARIMAX-GARCH

Com relação ao coeficiente R^2 , que foi usado para mensurar a quantidade de variação nas séries temporais que é explicada pelo método/modelo estimado, o modelo ARIMA-GARCH teve $R^2 = 0,361378$ enquanto o método

WARIMAX-GARCH teve $R^2 = 0,997249$. Estes resultados mostram que o método WARIMAX-GARCH explicou, aproximadamente, 99,72% da variação da série temporal de deslocamento do bloco F05, enquanto o modelo ARIMA-GARCH somente explicou 36,14%.

4.2 Resultados Computacionais: Método WARIMAX-GARCH NEURAL

4.2.1 Método WARIMAX-GARCH NEURAL: Série F05y

Neste experimento computacional foi utilizada a série F05y, descrita na Tabela 1 (página 60), com 2016 observações (dias) para treino e 30 observações (dias) foram utilizadas para teste, ou seja, horizonte de previsão $h = 30$, com 30 passos à frente (multipassos). Os valores aferidos pelo pêndulo automatizado, que compõe a série F05y, representam o deslocamento no sentido normal ao fluxo do rio. O gráfico da série temporal é apresentado na Figura 25.

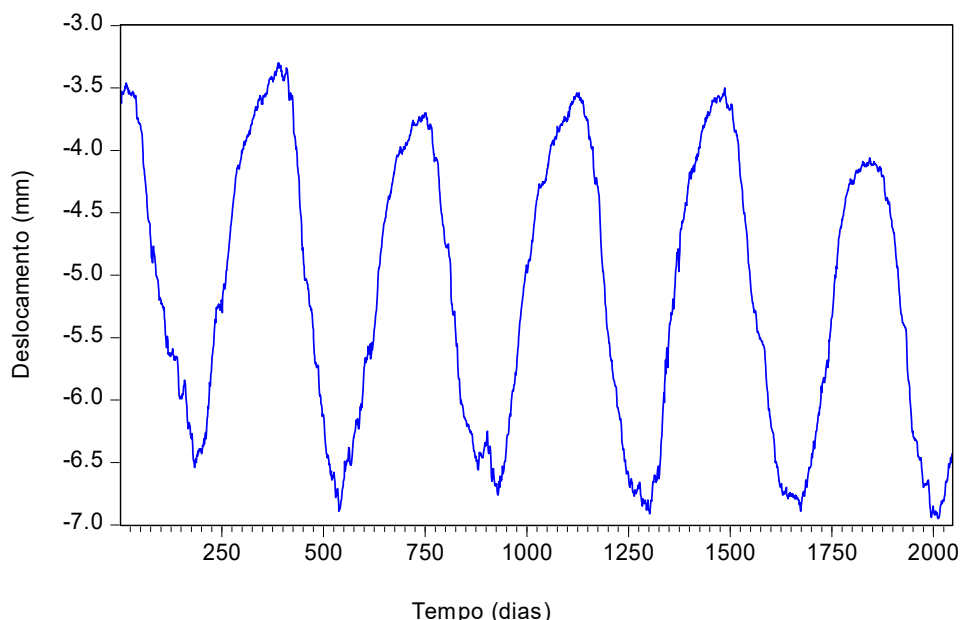


Figura 25 – Série temporal do deslocamento do bloco F05y

Nota-se, na Figura 25, que a série temporal apresenta oscilações de alta e baixa frequência nos dados do deslocamento. O gráfico da Figura 26, com os

90 últimos valores da série, evidencia mais explicitamente, as oscilações de alta frequência. O modelo WARIMAX-GARCH Neural pode modelar ambas as oscilações (de alta e baixa frequências) no domínio do tempo da série temporal subjacente, através dos modelos ARIMA-GARCH integrados com a abordagem de decomposição *Wavelet*.

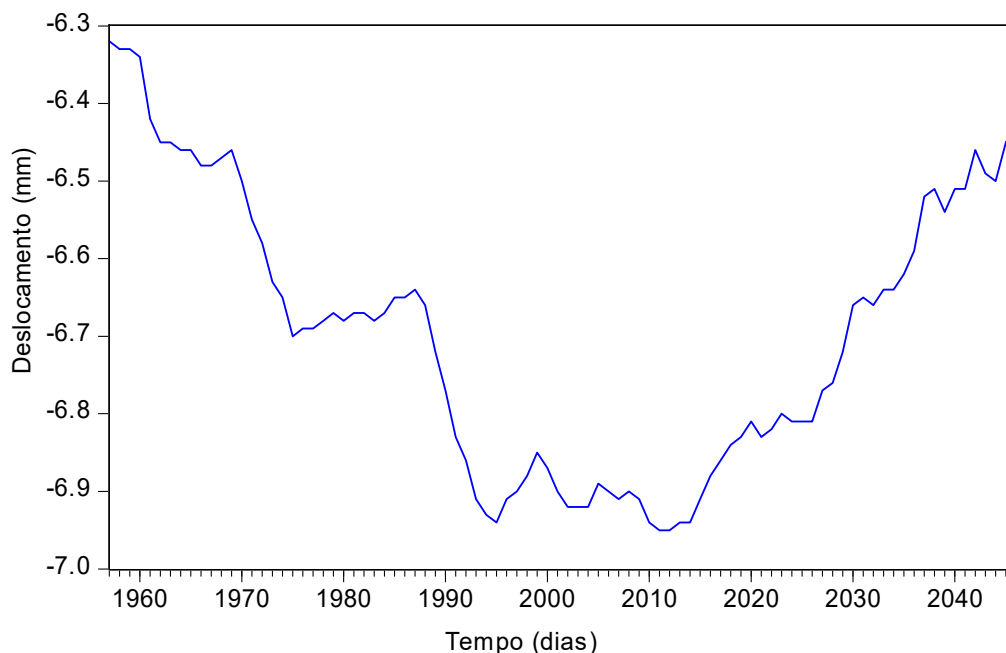


Figura 26 – Série temporal do deslocamento do bloco F05y

O teste da raiz unitária Dickey-Fuller Aumentado (DFA) foi aplicado e não houve rejeição, ao nível de significância de 1%, da hipótese de não estacionariedade dos dados de deslocamento. Porém, o teste DFA rejeitou fortemente a não estacionariedade da série temporal em sua segunda diferença (ver Apêndice II A). As funções de autocorrelação simples e parciais, em relação a este último, bem como sua versão ao quadrado, apresentam valores significativos ao nível de 1%, o que sugere a escolha de um modelo ARIMA-GARCH. Em função da indisponibilidade de variáveis exógenas convencionais apropriadas para serem utilizadas pelo modelo ARIMAX-GARCH para a série subjacente, foi apresentado um modelo ARIMA-GARCH que foi considerado plausível para o processo de modelagem.

Os modelos plausíveis obtidos nesta seção geraram previsões pontuais e intervalares utilizando a metodologia 1-passo dentro da amostra e multipassos fora da amostra, tanto para o modelo ARIMA-GARCH quanto para o método WARIMAX-GARCH Neural. Foi utilizado o software EViews para realizar os diagnósticos residuais para a validação estatística com o teste DFA, e obter os testes de Ljung-Box e Durbin-Watson para autocorrelação de primeira ordem, bem como os gráficos das funções de autocorreção e autocorrelação parciais.

4.2.1.1 Modelo ARIMA-GARCH

Um modelo ARIMA (3,1,1) integrado com o modelo GARCH (1,2) (isto é, modelo ARIMA-GARCH (3,1,1) x (1,2)), com distribuição gaussiana do erro, foi identificado para produzir o melhor ajuste para a amostra de treinamento y'_t ($t = 1, \dots, 2016$), uma vez diferenciada. Todos os parâmetros deste modelo foram estatisticamente significativos ao nível de 1%, e foram estimados por Máxima Verossimilhança.

Entre todos os modelos ARIMA-GARCH plausíveis, obtidos a partir da amostra de treinamento y'_t , foi o modelo ARIMA-GARCH (3,1,1) x (1,2) que produziu as melhores previsões para a média da série histórica subjacente, dentro da amostra de treinamento, em termos das estatísticas MAPE e MAE. Os resultados dos testes de Ljung-Box (Q-Stat) aplicados aos resíduos do modelo ARIMA-GARCH (3,1,1) x (1,2), conforme Figura 27 e Figura 28, sugerem que não há estrutura de autodependência linear significativa (ao nível de 1%) para os resíduos ordinários e quadráticos, até o $lag = 36$ (correspondendo a 36 dias). O teste para verificar o efeito ARCH também confirmou que não há autocorrelações significativas na variância residual (ao nível de 1%) para os lags de 1 até 36. Além disso, o cálculo da estatística de Durbin-Watson, com valor 1,967537, sugere que não há autocorrelação de $lag = 1$ significativa nos resíduos ordinários.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1	0.015	0.015	0.4690
		2	0.022	0.021	1.4021
		3	-0.018	-0.019	2.0710
		4	-0.042	-0.042	5.6430
		5	0.023	0.025	6.7327
		6	-0.005	-0.005	6.7898
		7	0.015	0.013	7.2668
		8	-0.009	-0.011	7.4443
		9	-0.002	-0.001	7.4547
		10	0.017	0.017	8.0080
		11	0.028	0.028	9.5659
		12	0.028	0.025	11.118
		13	0.021	0.021	12.047
		14	0.007	0.008	12.154
		15	-0.006	-0.004	12.223
		16	0.006	0.008	12.308
		17	0.026	0.026	13.633
		18	0.001	-0.001	13.635
		19	0.036	0.035	16.282
		20	-0.036	-0.035	18.881
		21	-0.041	-0.042	22.321
		22	-0.005	-0.006	22.379
		23	-0.002	-0.004	22.389
		24	-0.016	-0.022	22.944
		25	-0.018	-0.012	23.570
		26	0.025	0.022	24.813
		27	-0.010	-0.010	25.025
		28	0.036	0.030	27.617
		29	0.013	0.011	27.977
		30	0.010	0.009	28.188
		31	0.023	0.020	29.224
		32	0.018	0.020	29.909
		33	0.010	0.007	30.134
		34	0.018	0.018	30.779
		35	0.010	0.011	30.988
		36	-0.001	-0.000	30.989

Figura 27 – FAC e FACP dos resíduos ordinários padronizados

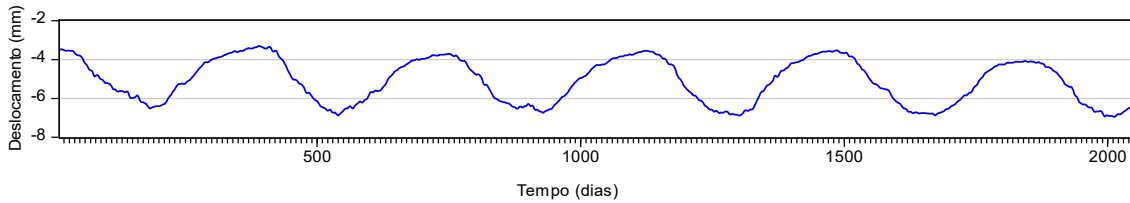
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1	0.061	0.061	7.5727
		2	-0.022	-0.026	8.5789
		3	-0.005	-0.002	8.6300
		4	-0.001	-0.001	8.6330
		5	-0.042	-0.042	12.245
		6	-0.000	0.005	12.246
		7	-0.018	-0.020	12.871
		8	-0.028	-0.026	14.411
		9	-0.038	-0.036	17.327
		10	-0.024	-0.023	18.463
		11	-0.051	-0.051	23.814
		12	-0.016	-0.014	24.358
		13	0.011	0.008	24.606
		14	0.009	0.003	24.763
		15	0.015	0.012	25.243
		16	0.027	0.020	26.766
		17	0.015	0.009	27.217
		18	-0.002	-0.005	27.224
		19	0.013	0.010	27.548
		20	-0.021	-0.026	28.460
		21	0.010	0.013	28.681
		22	0.019	0.016	29.411
		23	-0.001	-0.002	29.412
		24	-0.017	-0.011	29.986
		25	0.004	0.008	30.027
		26	0.027	0.030	31.505
		27	0.017	0.018	32.063
		28	0.005	0.006	32.109
		29	0.005	0.004	32.154
		30	0.061	0.064	39.877
		31	-0.017	-0.023	40.461
		32	-0.018	-0.012	41.140
		33	0.013	0.018	41.493
		34	0.028	0.028	43.094
		35	0.001	0.005	43.095
		36	0.017	0.020	43.684

Figura 28 – FAC e FACP dos resíduos quadráticos padronizados

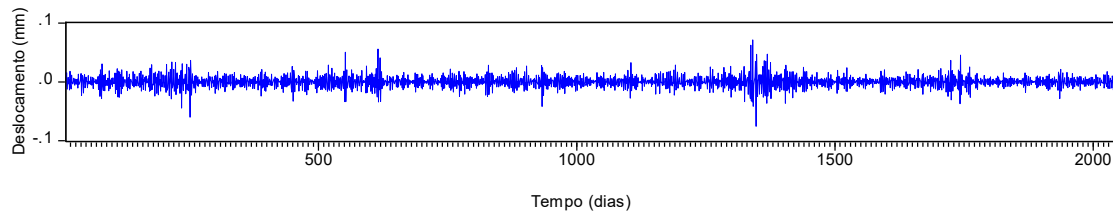
Vale ressaltar que a escolha do modelo ARIMA-GARCH, entre todos os modelos plausíveis identificados, foi determinada pela comparação do desempenho de previsão com o uso das estatísticas APE, MAPE, MAE e R^2 , em cada modelo candidato. O modelo selecionado foi aquele com menor APE, MAPE e MAE e com o maior R^2 .

4.2.1.2 Modelo WARIMAX-GARCH NEURAL

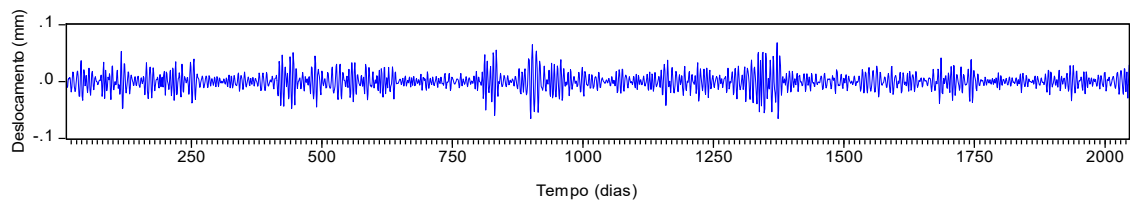
O método WARIMAX-GARCH Neural, conforme fluxograma da Figura 13, foi desenvolvido com base nos software MATLAB (versão 2013a) e EViews 8. Inicialmente, a decomposição *Wavelet* de nível $r = 2$ foi obtida a partir da amostra de treino da série de deslocamento F05y, no sentido normal a montante-jusante, do bloco F05 da barragem principal da Usina Hidrelétrica de Itaipu. Os gráficos das componentes *Wavelet*, com base ortonormal *db40* (DAUBECHIES, 1988), pode ser vista na Figura 29.



(a) Componente de aproximação de nível 2, $\tilde{y}_{A_2,t} (t = 1, \dots, 2016)$



(b) Componente de detalhe de nível 2, $\tilde{y}_{D_2,t} (t = 1, \dots, 2016)$



(c) Componente de detalhe de nível 3, $\tilde{y}_{D_3,t} (t = 1, \dots, 2016)$

Figura 29 – Componentes *Wavelet* da amostra de treino da série F05y

De acordo com a seção 2.1, para a decomposição *Wavelet* de nível $r = 2$, tem-se que $y_t = \tilde{y}_{A_2,t} + \tilde{y}_{D_2,t} + \tilde{y}_{D_3,t} + \varepsilon$, $(t = 1, \dots, 2016)$ onde ε é, estatisticamente, considerado ruído branco. Numa segunda etapa do método WARIMAX-GARCH Neural, cada uma das componentes *Wavelet* $\tilde{y}_{A_2,t}$, $\tilde{y}_{D_2,t}$ e $\tilde{y}_{D_3,t}$ $(t = 1, \dots, 2016)$ foi modelada, individualmente, por RNA's e por modelos ARIMA-GARCH apropriados para cada série. Estes modelos foram utilizados para gerar as previsões fora da amostra para suas respectivas médias, denotadas por $\hat{y}_{A_2,t}$, $\hat{y}_{D_2,t}$ e $\hat{y}_{D_3,t}$ $(t = 2017, \dots, 2046)$. O horizonte de previsão escolhido foi $h = 30$, com 30 passos à frente, ou seja, foram previstos 30 dias de deslocamento do bloco F05 (sentido normal ao fluxo do rio), dado pela série F05y. Todos os parâmetros dos três modelos estimados são estatisticamente significativos, a 1%, e o diagnóstico residual confirma sua plausibilidade (Ver Apêndice II D).

Numa terceira etapa, as três componentes *Wavelet* $\tilde{y}_{A_2,t,C}$, $\tilde{y}_{D_2,t,C}$ e $\tilde{y}_{D_3,t,C}$ (ou x_1, x_2 e x_3 , respectivamente) consistem de variáveis exógenas *Wavelet* geradas a partir da série de deslocamento F05y. Como mencionado anteriormente, elas são facilmente obtidas completando as componentes *Wavelet* com suas previsões fora da amostra, produzidas na segunda etapa.

Algebricamente, são representadas por:

- I. $x_{1,t} = \tilde{y}_{A_2,t,C} := \left((\tilde{y}_{A_2,t})_{t=1}^{2016}; (\hat{y}_{A_2,t})_{t=2017}^{2046} \right);$
- II. $x_{2,t} = \tilde{y}_{D_2,t,C} := \left((\tilde{y}_{D_2,t})_{t=1}^{2016}; (\hat{y}_{D_2,t})_{t=2017}^{2046} \right);$ e
- III. $x_{3,t} = \tilde{y}_{D_3,t,C} := \left((\tilde{y}_{D_3,t})_{t=1}^{2016}; (\hat{y}_{D_3,t})_{t=2017}^{2046} \right);$

Numa quarta etapa, foram usadas as variáveis exógenas *Wavelet* $x_{i,t}$ ($i = 1, 2, 3$) ($t = 1, \dots, 2046$) para o ajustamento do modelo WARIMAX-GARCH Neural utilizado para gerar as 30 previsões da série F05y (com 30 passos à frente). Algebricamente, a melhor formulação do modelo WARIMAX obtida foi:

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3) y_t = (1 - \theta_1 B) e_t + \sum_{i=1}^2 \tilde{y}_{A_2,t-i,C} \varphi_{A_2,t-i} + \tilde{y}_{D_2,t-1,C} \varphi_{D_2,t-1} + \tilde{y}_{D_3,t-2,C} \varphi_{D_3,t-2} + \varphi \sigma_t^2.$$

Também, o melhor modelo GARCH obtido foi GARCH (1,1), com distribuição do erro generalizada (GED), com a formulação algébrica que segue:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

O método da Máxima Verossimilhança foi utilizado para obter as melhores estimativas dos parâmetros WARIMAX-GARCH Neural. O Apêndice II D mostra que todos os parâmetros do modelo WARIMAX-GARCH foram considerados estatisticamente significativos ao nível de 1%.

A Figura 30 e a Figura 31 mostram, respectivamente, os gráficos da FAC e FACP (para lags de 1 até 36) dos resíduos ordinários padronizados e resíduos quadráticos padronizados associados ao modelo WARIMAX-GARCH Neural. Nota-se que todos os valores da FAC e FACP pertencem ao intervalo

de confiança com 99%, sugerindo que todos os valores são não significativos. A estatística de Ljung-Box (Q-Stat) sugere, também, que não há autodependência linear significativa, ao nível de 1% (dentro da amostra), para os resíduos ordinários e quadráticos padronizados provenientes do modelo WARIMAX-GARCH Neural. Também, o teste Ljung-Box, na Figura 31, confirma que não há efeito ARCH nos resíduos da previsão. A estatística de Durbin-Watson, cujo valor obtido foi 2,019100, confirmou a ausência de autocorrelação de primeira ordem, dentro da amostra, para os resíduos padronizados (ver Apêndice II E).

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 -0.008	-0.008	0.1260	
		2 0.029	0.029	1.8012	
		3 0.016	0.017	2.3482	
		4 -0.014	-0.015	2.7502	
		5 0.042	0.041	6.2862	0.012
		6 -0.009	-0.008	6.4681	0.039
		7 -0.003	-0.005	6.4816	0.090
		8 -0.036	-0.037	9.1208	0.058
		9 -0.040	-0.039	12.276	0.031
		10 -0.030	-0.031	14.107	0.028
		11 -0.015	-0.011	14.543	0.042
		12 0.009	0.011	14.717	0.065
		13 0.013	0.017	15.085	0.089
		14 0.014	0.017	15.503	0.115
		15 0.005	0.006	15.561	0.158
		16 0.000	-0.002	15.561	0.212
		17 0.020	0.015	16.355	0.231
		18 -0.014	-0.019	16.780	0.268
		19 0.026	0.021	18.138	0.255
		20 -0.049	-0.050	23.028	0.113
		21 0.043	0.044	26.794	0.061
		22 0.002	0.005	26.800	0.083
		23 0.009	0.014	26.968	0.105
		24 -0.016	-0.019	27.479	0.122
		25 -0.026	-0.020	28.814	0.119
		26 0.016	0.012	29.366	0.135
		27 -0.004	-0.001	29.403	0.167
		28 0.042	0.039	32.981	0.104
		29 0.021	0.022	33.871	0.111
		30 0.021	0.023	34.745	0.117
		31 0.030	0.029	36.596	0.103
		32 0.025	0.026	37.852	0.101
		33 0.008	0.002	37.982	0.123
		34 0.011	0.006	38.241	0.144
		35 0.008	0.005	38.360	0.170
		36 0.002	0.001	38.373	0.203

Figura 30 – FAC e FACP dos resíduos ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	0.029	0.029	1.7378	0.187	
2	-0.040	-0.041	4.9790	0.083	
3	0.022	0.025	5.9631	0.113	
4	0.006	0.003	6.0468	0.196	
5	0.000	0.002	6.0469	0.302	
6	-0.024	-0.024	7.1748	0.305	
7	0.007	0.009	7.2839	0.400	
8	-0.039	-0.042	10.407	0.238	
9	-0.035	-0.031	12.893	0.168	
10	-0.031	-0.032	14.782	0.140	
11	-0.016	-0.015	15.288	0.170	
12	-0.034	-0.035	17.596	0.129	
13	0.015	0.018	18.051	0.156	
14	0.026	0.022	19.465	0.148	
15	0.019	0.020	20.161	0.166	
16	0.028	0.026	21.726	0.152	
17	0.011	0.008	21.983	0.185	
18	0.001	-0.003	21.986	0.233	
19	0.073	0.071	32.687	0.026	
20	-0.020	-0.029	33.470	0.030	
21	-0.016	-0.010	33.983	0.036	
22	0.009	0.006	34.149	0.047	
23	0.010	0.012	34.346	0.060	
24	-0.009	-0.005	34.512	0.076	
25	0.003	0.013	34.529	0.097	
26	-0.010	-0.010	34.736	0.117	
27	0.021	0.030	35.595	0.124	
28	-0.007	-0.005	35.684	0.151	
29	-0.009	-0.005	35.846	0.178	
30	0.048	0.045	40.602	0.094	
31	-0.002	-0.001	40.607	0.116	
32	0.012	0.011	40.911	0.134	
33	0.024	0.020	42.101	0.133	
34	-0.017	-0.020	42.702	0.145	
35	0.029	0.031	44.394	0.133	
36	-0.002	-0.004	44.401	0.159	

Figura 31 – FAC e FACP dos resíduos quadráticos padronizados

Com base na Tabela 4, é possível comparar os resultados entre ARIMA-GARCH, RNA e WARIMAX-GARCH Neural dentro e fora da amostra de treino avaliando as estatísticas MAPE (*Mean Absolute Percent Error*), MAE (*Mean Absolute Error*) e RMSE (*Root Mean Squared Error*). Baseado nos três tipos de erro, o modelo WARIMAX-GARCH Neural teve melhor desempenho que o modelo ARIMA-GARCH (3,1,1) x (1,2) e a RNA (ver configurações da RNA no Apêndice II B), com base nas estatísticas MAPE, MAE e RMSE. Conforme Tabela 4, o método WARIMAX-GARCH Neural reduziu em até 91,1% o erro fora da amostra quando comparado ao método ARIMA-GARCH e em até 78,4% com relação à RNA. Isso significa que o WARIMAX-GARCH Neural modelou melhor a dinâmica da série e também teve melhor poder de gerar valores (previsões) fora da amostra.

MÉTODO	MAPE		MAE		RMSE	
	Dentro da amostra	Fora da amostra	Dentro da amostra	Fora da amostra	Dentro da amostra	Fora da amostra
ARIMA-GARCH	0,442%	2,931%	0,022	0,192	0,031	0,237
RNA	0,475%	1,232%	0,024	0,082	0,033	0,097
WARIMAX-GARCH	0,243%	0,280%	0,012	0,019	0,017	0,021

Neural						
--------	--	--	--	--	--	--

Tabela 4 – Valores dos erros MAPE, MAE e RMSE, dentro e fora da amostra, para a série temporal F05Y

A Figura 32 mostra o gráfico do Erro Percentual Absoluto (APE) calculado fora da amostra para cada passo à frente ($t = 2017, \dots, 2046$) para os modelos ARIMA-GARCH e WARIMAX-GARCH Neural. Nota-se que o modelo WARIMAX-GARCH Neural possui melhores valores APE que o modelo ARIMA-GARCH nos 30 valores fora da amostra, além de produzir previsões comparativamente melhores na medida em que o número de passos à frente vai aumentando. De fato, enquanto os erros de previsão do modelo ARIMA-GARCH mostram uma tendência positiva (crescente de pouco menos de 1% a quase 7%), o modelo WARIMAX-GARCH Neural mostra flutuações ao longo da linha de 1% com o aumento do horizonte de previsão. O Apêndice II E mostra outros gráficos da evolução temporal fora da amostra e as respectivas previsões de ambas as abordagens para apoiar a conclusão de que, de fato, o modelo WARIMAX-GARCH Neural tem melhor poder de generalização.

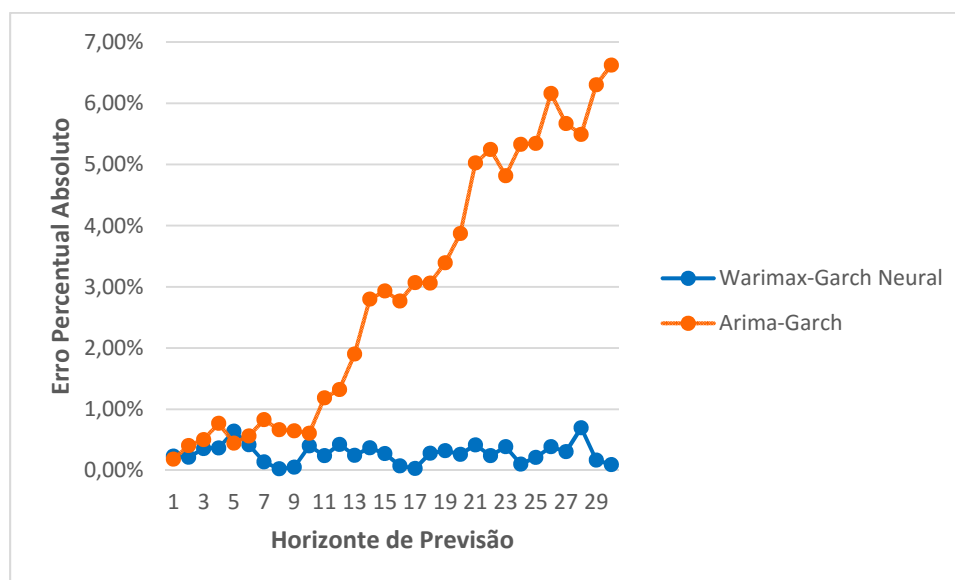


Figura 32 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH Neural e ARIMA-GARCH

A seguir são apresentados os gráficos dos deslocamentos reais e suas previsões, com intervalo de previsão de 99% para as previsões fora da amostra, obtidas por (a) Modelo ARIMA-GARCH (Figura 33) e (b) o Método WARIMAX-GARCH Neural (Figura 34). Nota-se, pela Figura 33, que as previsões obtidas pelo modelo ARIMA-GARCH foram maiores do que os valores observados em cada momento, no entanto, todos os valores permaneceram dentro do intervalo de 99% de confiança para as previsões fora da amostra.

Por outro lado, na Figura 34, as previsões obtidas pelo método WARIMAX-GARCH Neural descreveram a trajetória do deslocamento mais próxima do deslocamento real e, também, todos os valores permaneceram dentro do intervalo de 99% de confiança para a previsão, cujos limites foram muito menores que os do modelo ARIMA-GARCH. Isto é, a variância da densidade preditiva do modelo WARIMAX-GARCH Neural foi menor. Este gráfico também mostra a dinâmica das previsões produzidas pelo método WARIMAX-GARCH Neural por meio das oscilações das previsões em torno dos valores reais de deslocamento.

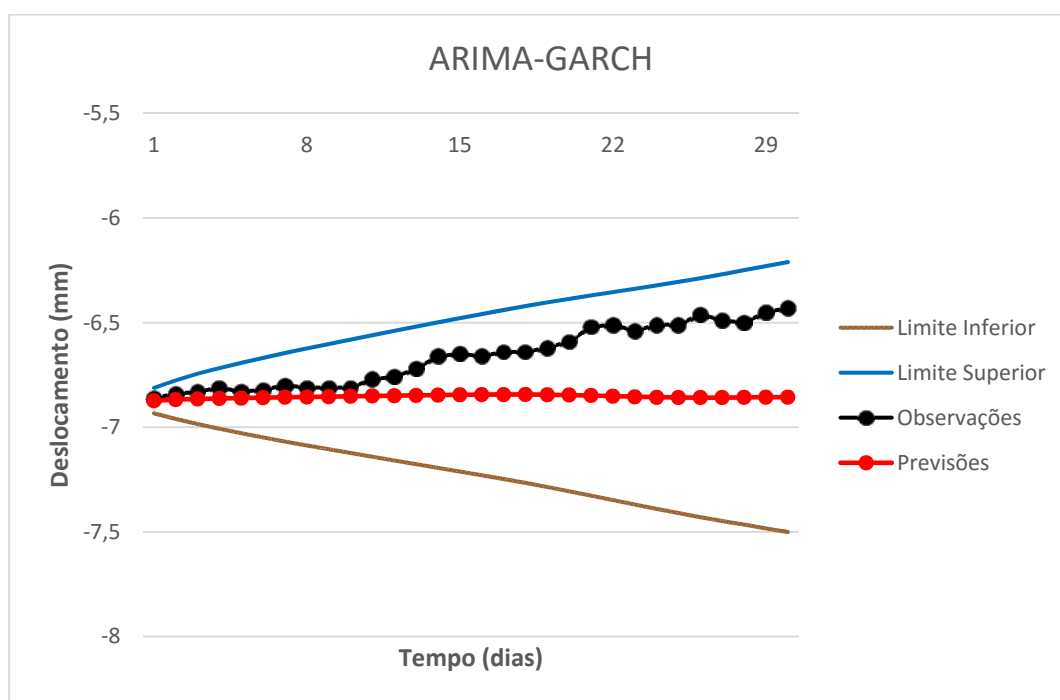


Figura 33 – Evolução temporal do intervalo de confiança para ARIMA-GARCH

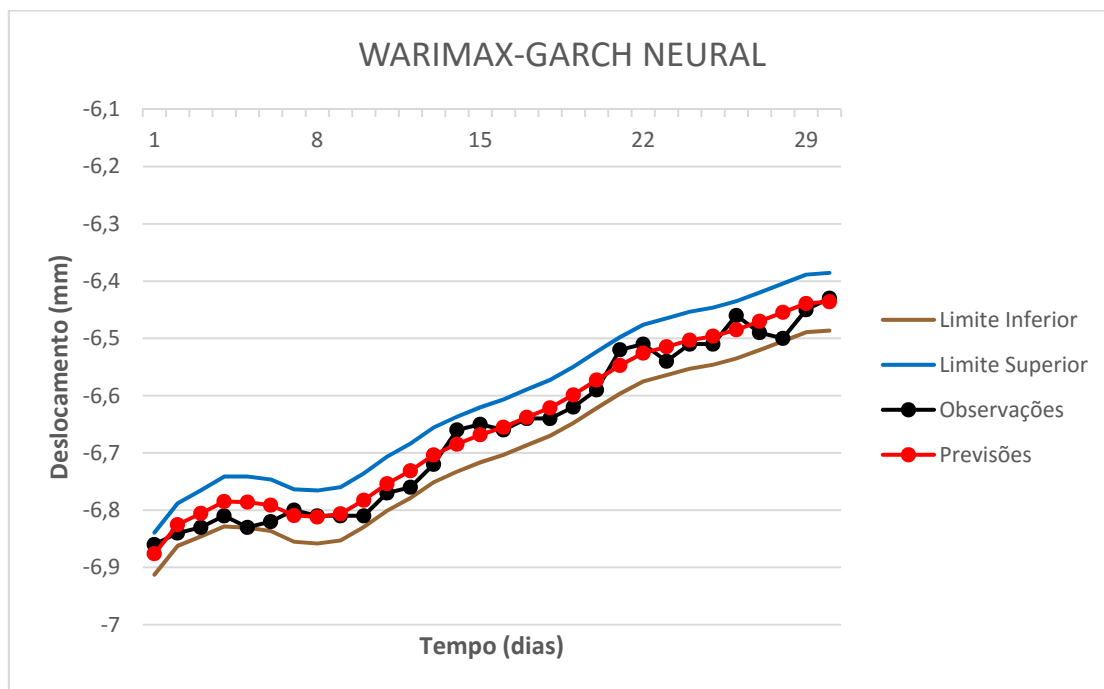


Figura 34 – Evolução Temporal do Intervalo de Confiança para WARIMAX-GARCH Neural

Com relação ao coeficiente R^2 , que foi usado para mensurar a quantidade de variação nas séries temporais que é explicada pelo método/modelo estimado, o método ARIMA-GARCH teve $R^2 = 0,265749$ enquanto o modelo WARIMAX-GARCH Neural teve $R^2 = 0,999784$. Estes resultados mostram que o método WARIMAX-GARCH Neural explicou, aproximadamente, 99,98% da variação da série temporal de deslocamento do bloco F05, no sentido normal ao fluxo do rio, enquanto o modelo ARIMA-GARCH somente explicou 26,57%.

4.2.2 Método WARIMAX-GARCH NEURAL: Série F13x

Neste experimento computacional foi utilizada a série F13x, descrita na Tabela 1 (página 60), com 1993 observações (dias) para treino e 30 observações (dias) foram utilizadas para teste, ou seja, horizonte de previsão $h = 30$, com 30 passos à frente (multipassos). Os valores aferidos pelo pêndulo automatizado, que compõe a série F13x, representam o deslocamento no

sentido normal ao fluxo do rio. O gráfico da série temporal é apresentado na Figura 35.

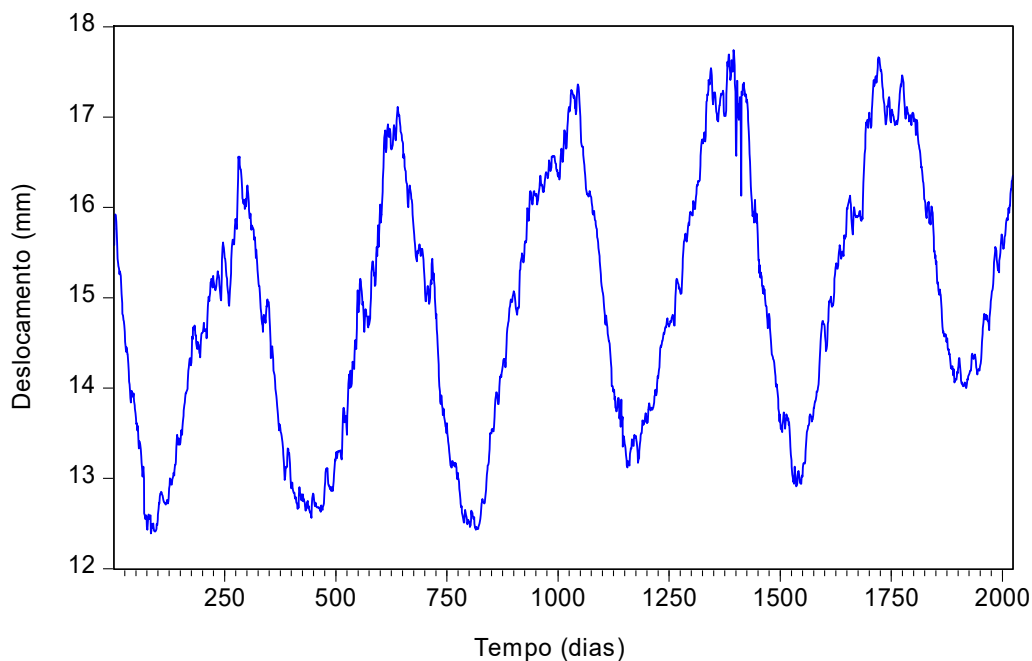


Figura 35 – Série temporal do deslocamento do bloco F13x

A partir da Figura 35 é possível observar que a série temporal apresenta oscilações de alta e baixa frequência nos dados do deslocamento. O gráfico da Figura 36, com os 90 últimos valores da série, evidencia mais explicitamente, as oscilações de alta frequência. Um modelo WARIMAX-GARCH Neural pode modelar ambas as variações no domínio do tempo da série temporal subjacente, através dos modelos ARIMA-GARCH integrados com a abordagem de decomposição *Wavelet*.

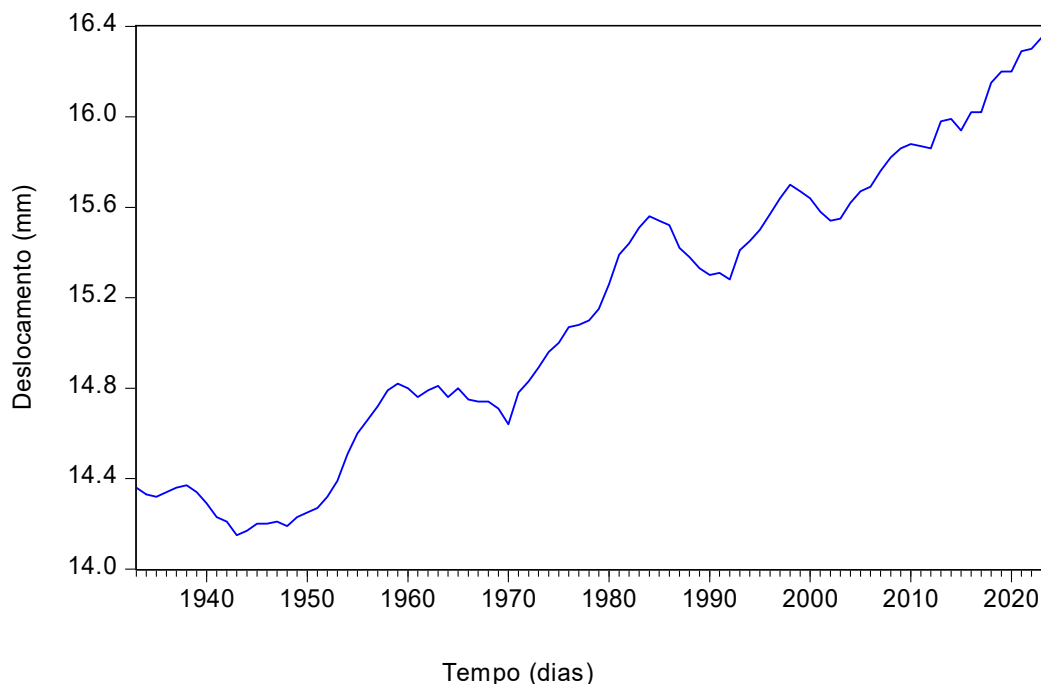


Figura 36 – Série temporal do deslocamento do bloco F13x

O teste da raiz unitária Dickey-Fuller Aumentado (DFA) foi aplicado e não houve rejeição, ao nível de significância de 1%, da hipótese de não-estacionariedade dos dados de deslocamento. Porém, o teste DFA rejeitou fortemente a não estacionariedade da série temporal em sua segunda diferença (ver Apêndice III A). As funções de autocorrelação simples e parciais, em relação a este último, bem como sua versão ao quadrado, apresentam valores significativos ao nível de 1%, o que sugere a escolha de um modelo ARIMA-GARCH. Em função da indisponibilidade de variáveis exógenas convencionais apropriadas para serem utilizadas pelo modelo ARIMAX-GARCH para a série subjacente, foi apresentado um modelo ARIMA-GARCH que é considerado plausível para o processo de modelagem.

Os modelos plausíveis obtidos nesta seção geraram previsões pontuais e intervalares utilizando a metodologia 1-passo dentro da amostra e multipassos fora da amostra, tanto para o modelo ARIMA-GARCH quanto para o método WARIMAX-GARCH Neural. Foi utilizado o software EViews para realizar os diagnósticos residuais para a validação estatística com o teste DFA, e obter os

testes de Ljung-Box e Durbin-Watson para autocorrelação de primeira ordem, bem como os gráficos das funções de autocorreção e autocorrelação parciais.

4.2.2.1 Modelo ARIMA-GARCH

Um modelo ARIMA (2,1,18) integrado com o modelo GARCH (1,1) (isto é, modelo ARIMA-GARCH (2,1,18) x (1,1)), com distribuição gaussiana do erro, foi identificado para produzir o melhor ajuste para a amostra de treinamento y'_t ($t = 1, \dots, 1993$), uma vez diferenciada. Todos os parâmetros deste modelo foram estatisticamente significativos ao nível de 1%, e foram estimados por Máxima Verossimilhança.

Entre todos os modelos ARIMA-GARCH plausíveis, obtidos a partir da amostra de treinamento y'_t , foi o modelo ARIMA-GARCH (2,1,18) x (1,1) que produziu as melhores previsões para a média da série histórica subjacente, dentro da amostra de treinamento, em termos das estatísticas MAPE e MAE. Os resultados dos testes de Ljung-Box (Q-Stat) aplicados aos resíduos ordinários do modelo ARIMA-GARCH (2,1,18) x (1,1), conforme Figura 37 e Figura 38, sugerem que não há estrutura de autodependência linear significativa (ao nível de 1%) para os resíduos ordinários e quadráticos, até o $lag = 36$. A Figura 38 exibe a ausência de autocorrelações significativas para os resíduos quadráticos (ao nível de 1%) para os lags de 1 até 36. Além disso, o cálculo da estatística de Durbin-Watson, com valor 2,325354, sugere que não há autocorrelação de $lag = 1$ significativa nos resíduos ordinários.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1	-0.006	-0.006	0.0764
		2	0.019	0.019	0.7846
		3	-0.008	-0.008	0.9234
		4	-0.035	-0.035	3.3450
		5	-0.008	-0.008	3.4792
		6	0.005	0.006	3.5240
		7	0.033	0.032	5.6513
		8	-0.030	-0.032	7.4988
		9	-0.004	-0.006	7.5294
		10	-0.004	-0.002	7.5635
		11	0.007	0.009	7.6548
		12	0.048	0.047	12.289
		13	-0.020	-0.021	13.076
		14	0.036	0.033	15.686
		15	0.016	0.021	16.207
		16	-0.011	-0.010	16.469
		17	0.009	0.008	16.626
		18	0.010	0.012	16.832
		19	0.018	0.017	17.464
		20	-0.023	-0.019	18.488
		21	0.019	0.016	19.236
		22	0.011	0.014	19.459
		23	-0.033	-0.031	21.608
		24	0.006	0.001	21.675
		25	-0.042	-0.039	25.258
		26	0.029	0.026	26.976
		27	0.003	0.005	26.993
		28	0.033	0.029	29.208
		29	0.025	0.021	30.431
		30	0.008	0.011	30.570
		31	0.025	0.023	31.802
		32	-0.023	-0.018	32.882
		33	0.014	0.007	33.256
		34	0.013	0.017	33.578
		35	-0.018	-0.016	34.246
		36	-0.025	-0.030	35.566

Figura 37 – FAC e FACP dos resíduos ordinários padronizados

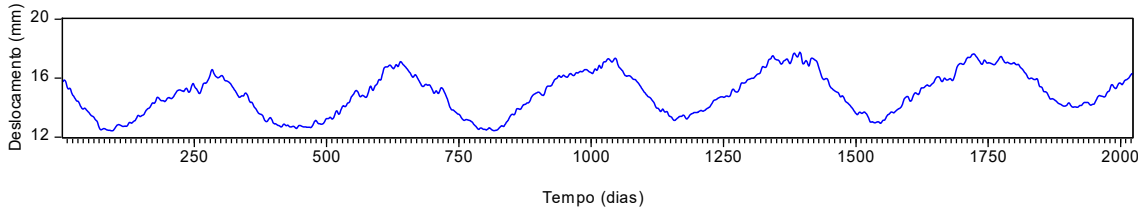
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1	0.027	0.027	1.4287
		2	-0.009	-0.010	1.6085
		3	-0.023	-0.023	2.6849
		4	-0.019	-0.018	3.4378
		5	-0.010	-0.009	3.6385
		6	-0.000	-0.001	3.6386
		7	-0.027	-0.028	5.1151
		8	0.009	0.009	5.2669
		9	-0.015	-0.016	5.6974
		10	-0.013	-0.013	6.0247
		11	0.004	0.004	6.0564
		12	0.047	0.046	10.520
		13	0.090	0.087	26.893
		14	-0.012	-0.017	27.188
		15	0.012	0.017	27.478
		16	0.047	0.051	31.850
		17	0.018	0.019	32.473
		18	0.005	0.007	32.522
		19	0.024	0.029	33.708
		20	0.010	0.016	33.896
		21	0.045	0.047	37.971
		22	-0.012	-0.007	38.251
		23	-0.019	-0.012	38.974
		24	-0.011	-0.010	39.215
		25	-0.009	-0.014	39.396
		26	-0.003	-0.006	39.409
		27	0.029	0.030	41.111
		28	0.006	-0.002	41.173
		29	0.008	-0.003	41.300
		30	0.009	0.007	41.448
		31	-0.011	-0.014	41.694
		32	0.010	0.002	41.907
		33	0.019	0.011	42.662
		34	0.017	0.009	43.274
		35	0.008	0.009	43.405
		36	-0.001	0.001	43.406

Figura 38 – FAC e FACP dos resíduos quadráticos padronizados

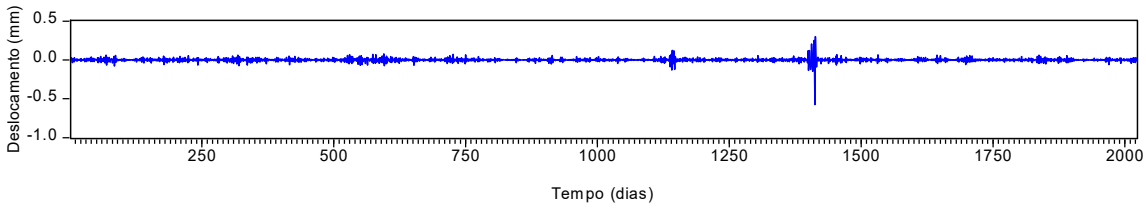
Vale ressaltar que a escolha do modelo ARIMA-GARCH, entre todos os modelos plausíveis identificados, foi determinada pela comparação do desempenho de previsão com o uso das estatísticas APE, MAPE, MAE e R^2 , em cada modelo candidato. O modelo selecionado foi aquele com menor APE, MAPE e MAE e com o maior R^2 .

4.2.2.2 Modelo WARIMAX-GARCH NEURAL

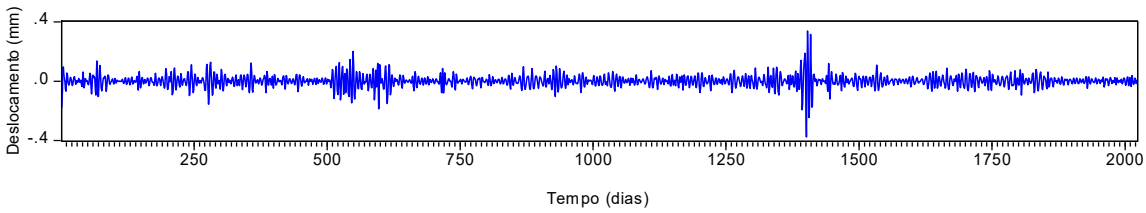
O método WARIMAX-GARCH Neural, conforme fluxograma da Figura 13, foi desenvolvido com base nos software MATLAB (versão 2013a) e EViews 8. Inicialmente, a decomposição *Wavelet* de nível $r = 2$ foi obtida a partir da amostra de treino da série de deslocamento, no sentido montante-jusante, do bloco F13 da barragem principal da Usina Hidrelétrica de Itaipu. Os gráficos das componentes *Wavelet*, com base ortonormal *Coiflet* de momento nulo 5, podem ser vistos na Figura 39.



(a) Componente de aproximação de nível 2, $\tilde{y}_{A_2,t}(t = 1, \dots, 1993)$



(b) Componente de detalhe de nível 2, $\tilde{y}_{D_2,t}(t = 1, \dots, 1993)$



(c) Componente de detalhe de nível 3, $\tilde{y}_{D_3,t}(t = 1, \dots, 1993)$

Figura 39 – Componentes *Wavelet* da amostra de treino da série F13x

De acordo com a seção 2.1, para a decomposição *Wavelet* de nível $r = 2$, tem-se que $y_t = \tilde{y}_{A_2,t} + \tilde{y}_{D_2,t} + \tilde{y}_{D_3,t} + \varepsilon$, $(t = 1, \dots, 1993)$ onde ε é, estatisticamente, considerado ruído branco. Na segunda etapa do método WARIMAX-GARCH Neural, cada uma das componentes *Wavelet* $\tilde{y}_{A_2,t}$, $\tilde{y}_{D_2,t}$ e $\tilde{y}_{D_3,t}$ ($t = 1, \dots, 1993$) foi modelada, individualmente, por RNA's e por modelos ARIMA-GARCH apropriados para cada série. Estes modelos foram utilizados para gerar as previsões fora da amostra para suas respectivas médias, denotadas por $\hat{y}_{A_2,t}$, $\hat{y}_{D_2,t}$ e $\hat{y}_{D_3,t}$ ($t = 1994, \dots, 2023$). O horizonte de previsão escolhido foi $h = 30$, com 30 passos à frente, ou seja, foram previstos 30 dias de deslocamento do bloco F13 (no sentido do fluxo do rio), dada pela série F13x. Todos os parâmetros dos três modelos estimados são estatisticamente significativos, a 1%, e o diagnóstico residual confirma sua plausibilidade (Ver Apêndice III D).

Numa terceira etapa, as três componentes *Wavelet* $\tilde{y}_{A_2,t,C}$, $\tilde{y}_{D_2,t,C}$ e $\tilde{y}_{D_3,t,C}$ (ou x_1, x_2 e x_3 , respectivamente) consistem de variáveis exógenas *Wavelet* geradas a partir da série de deslocamento F13x. Como mencionado anteriormente, elas são facilmente obtidas completando as componentes *Wavelet* com suas previsões fora da amostra, produzidas na segunda etapa.

Algebricamente, são representadas por:

$$\text{IV. } x_{1,t} = \tilde{y}_{A_2,t,C} := \left((\tilde{y}_{A_2,t})_{t=1}^{1993}; (\hat{y}_{A_2,t})_{t=1994}^{2023} \right);$$

$$\text{V. } x_{2,t} = \tilde{y}_{D_2,t,C} := \left((\tilde{y}_{D_2,t})_{t=1}^{1993}; (\hat{y}_{D_2,t})_{t=1994}^{2023} \right); e$$

$$\text{VI. } x_{3,t} = \tilde{y}_{D_3,t,C} := \left((\tilde{y}_{D_3,t})_{t=1}^{1993}; (\hat{y}_{D_3,t})_{t=1994}^{2023} \right);$$

Numa quarta etapa, foram usadas as variáveis exógenas *Wavelet* $x_{i,t}$ ($i = 1, 2, 3$) ($t = 1, \dots, 1993$) para o ajustamento do modelo WARIMAX-GARCH Neural utilizado para gerar as 30 previsões da série F13x (com 30 passos à frente). Algebricamente, a melhor formulação do modelo WARIMAX obtida foi:

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3) y_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) e_t + \sum_{i=1}^2 \tilde{y}_{A_2,t-i,C} \varphi_{A_2,t-i} + \tilde{y}_{D_2,t-1,C} \varphi_{D_2,C,t-1} + \tilde{y}_{D_3,t-2,C} \varphi_{D_3,C,t-2} + \varphi \sigma_t^2.$$

Também, o melhor modelo GARCH obtido foi GARCH (1,1), com distribuição do erro generalizada (GED), com a formulação algébrica que segue:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

O método da Máxima Verossimilhança foi utilizado para obter as melhores estimativas dos parâmetros WARIMAX-GARCH. O Apêndice III D mostra que todos os parâmetros do modelo WARIMAX-GARCH foram considerados estatisticamente significativos ao nível de 1%.

O Apêndice III E exibe os gráficos da FAC e FACP (para lags de 1 até 36) dos resíduos ordinários padronizados e resíduos quadráticos padronizados associados ao modelo WARIMAX-GARCH Neural. Nota-se que todos os valores da FAC e FACP pertencem ao intervalo de confiança com 99%, sugerindo que todos os valores são não significativos. O teste de Ljung-Box (Q-

Stat) sugere, também, que não há autodependência linear significativa, ao nível de 1% (dentro da amostra), para os resíduos ordinários e quadráticos padronizados provenientes do modelo WARIMAX-GARCH Neural. A estatística de Durbin-Watson, cujo valor obtido foi 2,092066, confirmou a ausência de autocorrelação de primeira ordem, dentro da amostra, para os resíduos padronizados.

Com base na Tabela 5, é possível comparar os resultados entre ARIMA-GARCH, RNA e WARIMAX-GARCH Neural dentro e fora da amostra de treino avaliando as estatísticas MAPE (*Mean Absolute Percent Error*) e MAE (*Mean Absolute Error*). Baseado nos dois tipos de erro, o modelo WARIMAX-GARCH Neural teve melhor desempenho que o modelo ARIMA-GARCH (2,1,18) x (1,1) e a RNA (ver configurações da RNA no Apêndice III C), com base nas estatísticas MAPE, MAE e RMSE. Isso significa que o WARIMAX-GARCH Neural modelou melhor a dinâmica da série e também teve melhor poder de gerar valores (previsões) fora da amostra. A redução do erro obtido, fora da amostra, pelo método WARIMAX-GARCH Neural em relação ao modelo ARIMA-GARCH foi de até 74,1% e, comparado à RNA foi de até 50,4%, conforme Tabela 5.

MÉTODO	MAPE		MAE		RMSE	
	<i>Dentro da amostra</i>	<i>Fora da amostra</i>	<i>Dentro da amostra</i>	<i>Fora da amostra</i>	<i>Dentro da amostra</i>	<i>Fora da amostra</i>
ARIMA-GARCH	0,325%	0,794%	0,00487	0,1272	0,077	0,166
RNA	0,346%	0,438%	0,0518	0,0697	0,0773	0,0868
WARIMAX-GARCH Neural	0,152%	0,220%	0,0227	0,0349	0,036	0,043

Tabela 5 – Valores dos erros MAPE, MAE e RMSE, dentro e fora da amostra, para a série temporal F13X

A Figura 40 mostra o gráfico do Erro Percentual Absoluto (APE) calculado fora da amostra para cada passo à frente ($t = 1994, \dots, 2023$) para os modelos ARIMA-GARCH e WARIMAX-GARCH Neural. Nota-se que o modelo WARIMAX-GARCH Neural possui melhores valores APE que o modelo ARIMA-

GARCH nos 30 valores fora da amostra, além de produzir previsões comparativamente melhores na medida em que o número de passos à frente vai aumentando. De fato, enquanto os erros de previsão do modelo ARIMA-GARCH mostram uma tendência positiva a partir do horizonte de previsão $h = 19$ (crescente de pouco menos de 1% a quase 2,5%), o modelo WARIMAX-GARCH Neural mostra flutuações abaixo da linha de 0,5% com o aumento do horizonte de previsão. O Apêndice III E mostra outros gráficos da evolução temporal fora da amostra e as respectivas previsões de ambas as abordagens para apoiar a conclusão de que, de fato, o modelo WARIMAX-GARCH Neural tem melhor poder de generalização.

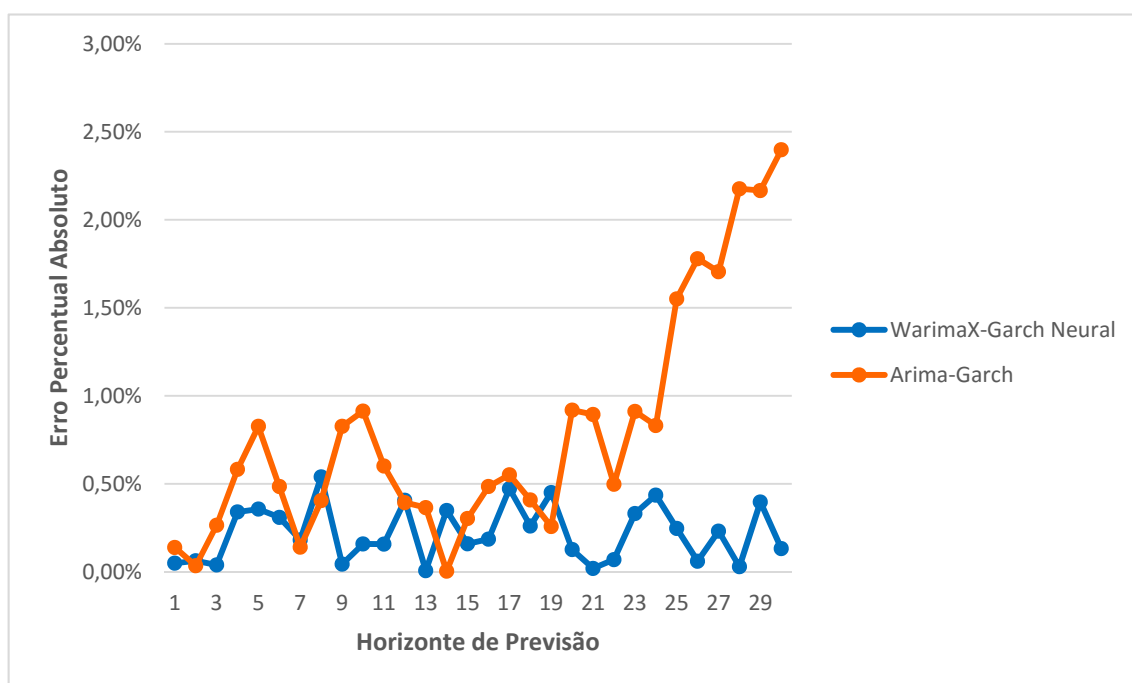


Figura 40 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH Neural e ARIMA-GARCH

São apresentados, a seguir, os gráficos dos deslocamentos reais e suas previsões, com intervalo de previsão de 99% para as previsões fora da amostra, obtidas por (a) Modelo ARIMA-GARCH (Figura 41) e (b) o Método WARIMAX-GARCH Neural (Figura 42). Nota-se, pela Figura 41, que as previsões obtidas pelo modelo ARIMA-GARCH foram maiores do que os

valores observados em cada momento, no entanto, todos os valores permaneceram dentro do intervalo de 99% de confiança para as previsões fora da amostra.

Por outro lado, na Figura 42, as previsões obtidas pelo método WARIMAX-GARCH Neural descreveram a trajetória do deslocamento mais próxima do deslocamento real e, também, todos os valores permaneceram dentro do intervalo de 99% de confiança para a previsão, cujos limites foram muito menores que os do modelo ARIMA-GARCH. Isto é, a variância da densidade preditiva do modelo WARIMAX-GARCH Neural foi menor. Este gráfico também mostra a dinâmica das previsões produzidas pelo método WARIMAX-GARCH Neural por meio das oscilações das previsões em torno dos valores reais de deslocamento.

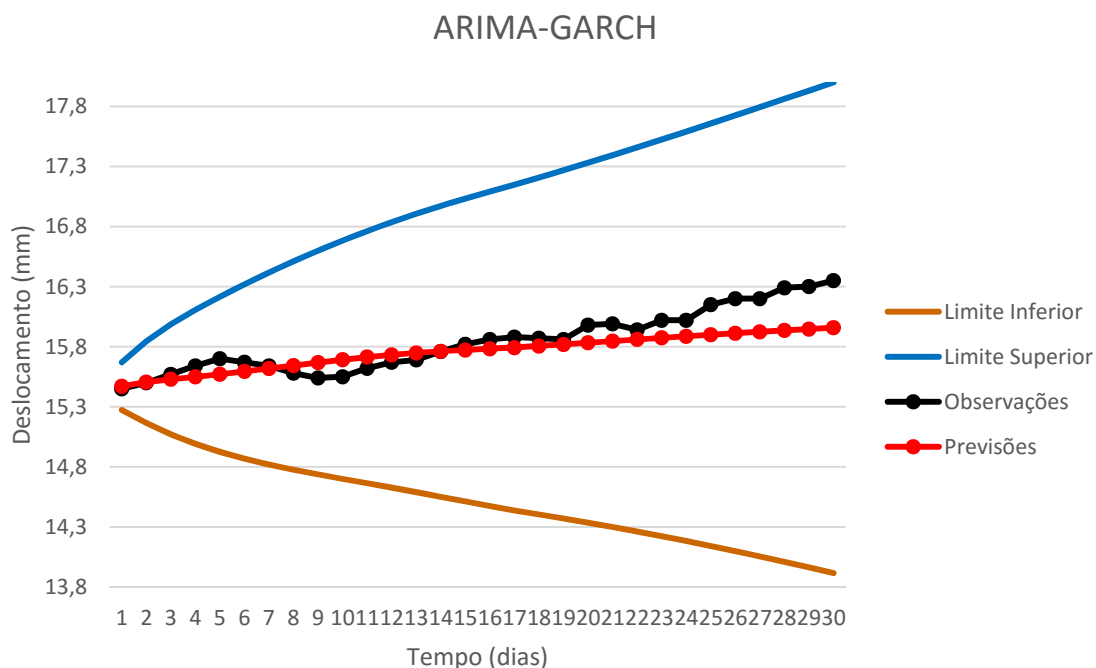


Figura 41 – Evolução temporal do intervalo de confiança para ARIMA-GARCH

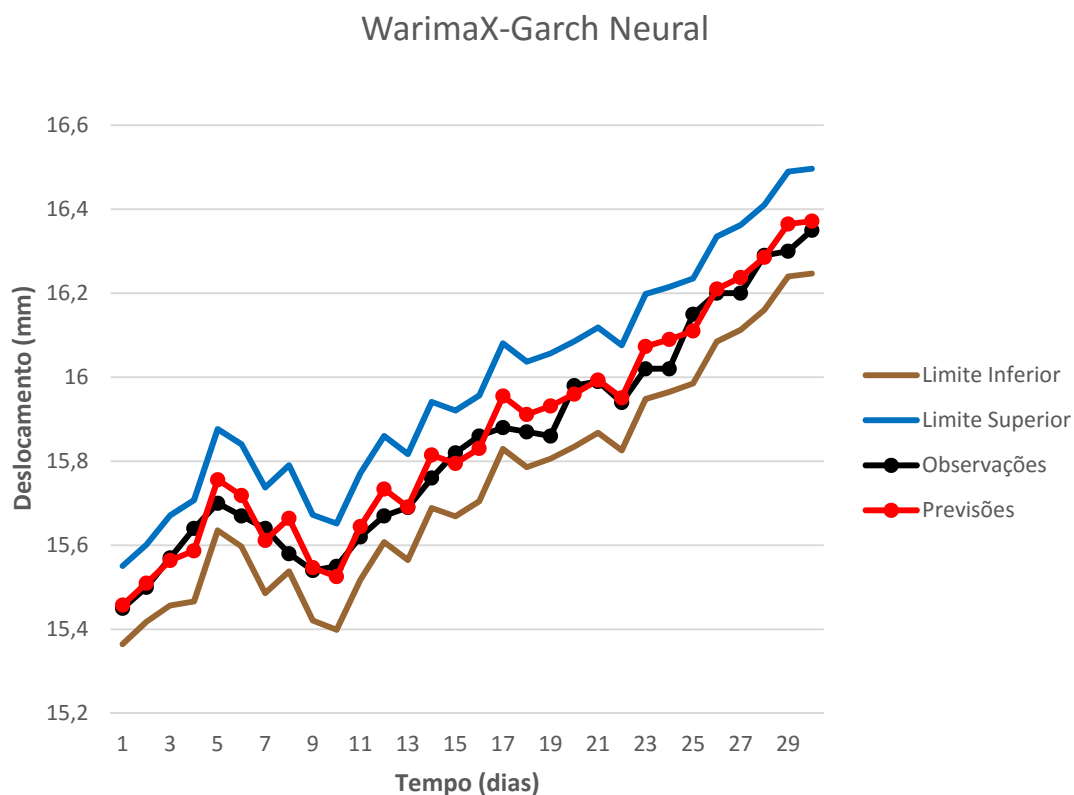


Figura 42 – Evolução Temporal do Intervalo de Confiança para WARIMAX-GARCH Neural

Com relação ao coeficiente de R^2 , que foi usado para mensurar a quantidade de variação nas séries temporais que é explicada pelo método/modelo estimado, o modelo ARIMA-GARCH teve $R^2 = 0,074383$ enquanto o método WARIMAX-GARCH Neural teve $R^2 = 0,999384$. Estes resultados mostram que o método WARIMAX-GARCH Neural explicou, aproximadamente, 99,94% da variação da série temporal de deslocamento do bloco F13, no sentido normal ao fluxo do rio, enquanto o modelo ARIMA-GARCH somente explicou 7,44%.

4.2.3 Método WARIMAX-GARCH NEURAL: Série F13y

Para a série F13y (ver descrição na Tabela 1, página 60) foram utilizadas 1993 observações (dias) para treinamento do modelo e 30 observações (dias)

para teste, ou seja, horizonte de previsão $h = 30$, com 30 passos à frente (multipassos). Os valores aferidos pelo pêndulo automatizado, que compõe a série F13y, representam o deslocamento do bloco F13, da barragem principal da Usina de Itaipu, no sentido do fluxo do rio. O gráfico desta série temporal é apresentado na Figura 43.

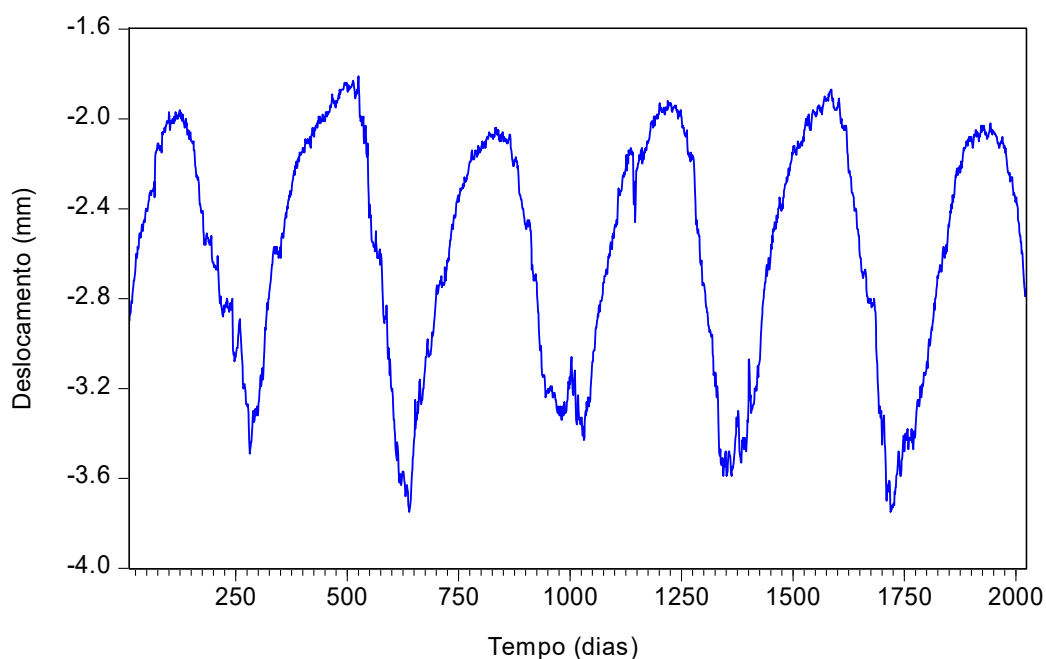


Figura 43 – Série temporal do deslocamento do bloco F13y

Na Figura 43 é possível observar que a série temporal apresenta oscilações de alta e baixa frequência nos dados do deslocamento. O gráfico da Figura 44, com os 90 últimos valores da série, evidencia mais explicitamente, as oscilações de alta frequência. Um modelo WARIMAX-GARCH Neural pode modelar ambas as variações no domínio do tempo da série temporal subjacente, através dos modelos ARIMA-GARCH integrados com a abordagem de decomposição *Wavelet*.

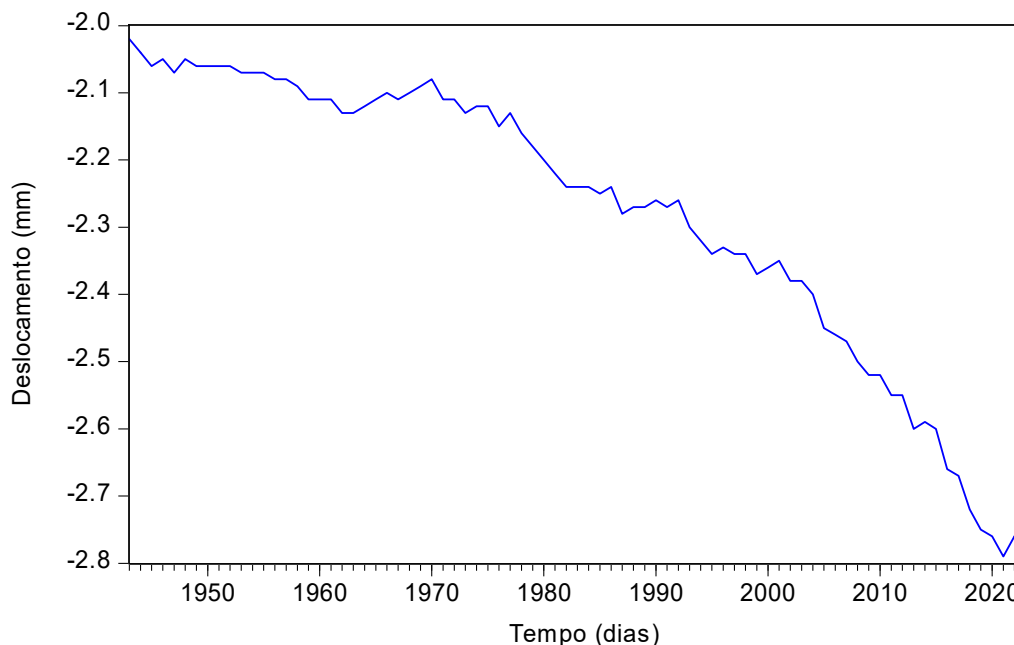


Figura 44 – Série temporal do deslocamento do bloco F13y

O teste da raiz unitária Dickey-Fuller Aumentado (DFA) foi aplicado e não houve rejeição, ao nível de significância de 1%, da hipótese de não-estacionariedade dos dados de deslocamento. Porém, o teste DFA rejeitou fortemente a não estacionariedade da série temporal em sua segunda diferença (ver Apêndice IV A). As funções de autocorrelação simples e parciais, em relação a este último, bem como sua versão ao quadrado, apresentam valores significativos ao nível de 1%, o que sugere a escolha de um modelo ARIMA-GARCH. O modelo ARIMA-GARCH foi considerado plausível para o processo de modelagem, uma vez que não houve disponibilidade de variáveis exógenas convencionais apropriadas para serem utilizadas pelo modelo ARIMAX-GARCH para a série subjacente.

Os modelos plausíveis obtidos nesta seção geraram previsões pontuais e intervalares utilizando a metodologia 1-passo dentro da amostra e multipassos fora da amostra, tanto para o modelo ARIMA-GARCH quanto para o método WARIMAX-GARCH Neural. Foi utilizado o software EViews para realizar os diagnósticos residuais para a validação estatística com o teste DFA, e obter os testes de Ljung-Box e Durbin-Watson para autocorrelação de primeira ordem, bem como os gráficos das funções de autocorreção e autocorrelação parciais.

4.2.3.1 Modelo ARIMA-GARCH

Um modelo ARIMA (5,0,3) integrado com o modelo GARCH (2,1) (isto é, modelo ARIMA-GARCH (5,0,3) x (2,1)), erro com distribuição t de *Student*, foi identificado para produzir o melhor ajuste para a amostra de treinamento y'_t ($t = 1, \dots, 1993$). Todos os parâmetros deste modelo foram estatisticamente significativos ao nível de 1%, e foram estimados por Máxima Verossimilhança.

Entre todos os modelos ARIMA-GARCH plausíveis, obtidos a partir da amostra de treinamento y'_t , foi o modelo ARIMA-GARCH (5,0,3) x (2,1) que produziu as melhores previsões para a média da série histórica subjacente, dentro da amostra de treinamento, em termos das estatísticas MAPE e MAE. Os resultados dos testes de Ljung-Box (Q-Stat) aplicados aos resíduos do modelo ARIMA-GARCH (5,0,3) x (2,1), conforme Figura 45 e Figura 46, sugerem que não há estrutura de autodependência linear significativa (ao nível de 1%) para os resíduos ordinários e quadráticos, até o $lag = 36$ (correspondendo a 36 dias). A Figura 46 mostra a ausência de autocorrelações significativas nos erros quadráticos padronizados (ao nível de 1%) para os lags de 1 até 36 (Ver Apêndice IV C). Além disso, o cálculo da estatística de Durbin-Watson, com valor 1,746919, sugere que não há autocorrelação de $lag = 1$ significativa nos resíduos ordinários.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 -0.006	-0.006	0.0764	
		2 0.019	0.019	0.7846	
		3 -0.008	-0.008	0.9234	
		4 -0.035	-0.035	3.3450	
		5 -0.008	-0.008	3.4792	0.062
		6 0.005	0.006	3.5240	0.172
		7 0.033	0.032	5.6513	0.130
		8 -0.030	-0.032	7.4988	0.112
		9 -0.004	-0.006	7.5294	0.184
		10 -0.004	-0.002	7.5635	0.272
		11 0.007	0.009	7.6548	0.364
		12 0.048	0.047	12.289	0.139
		13 -0.020	-0.021	13.076	0.159
		14 0.036	0.033	15.686	0.109
		15 0.016	0.021	16.207	0.134
		16 -0.011	-0.010	16.469	0.171
		17 0.009	0.008	16.626	0.217
		18 0.010	0.012	16.832	0.265
		19 0.018	0.017	17.464	0.292
		20 -0.023	-0.019	18.488	0.296
		21 0.019	0.016	19.236	0.315
		22 0.011	0.014	19.459	0.364
		23 -0.033	-0.031	21.608	0.304
		24 0.006	0.001	21.675	0.358
		25 -0.042	-0.039	25.258	0.236
		26 0.029	0.026	26.976	0.212
		27 0.003	0.005	26.993	0.256
		28 0.033	0.029	29.208	0.212
		29 0.025	0.021	30.431	0.209
		30 0.008	0.011	30.570	0.245
		31 0.025	0.023	31.802	0.240
		32 -0.023	-0.018	32.882	0.240
		33 0.014	0.007	33.256	0.268
		34 0.013	0.017	33.578	0.298
		35 -0.018	-0.016	34.246	0.315
		36 -0.025	-0.030	35.566	0.304

Figura 45 – FAC e FACP dos resíduos ordinários padronizados

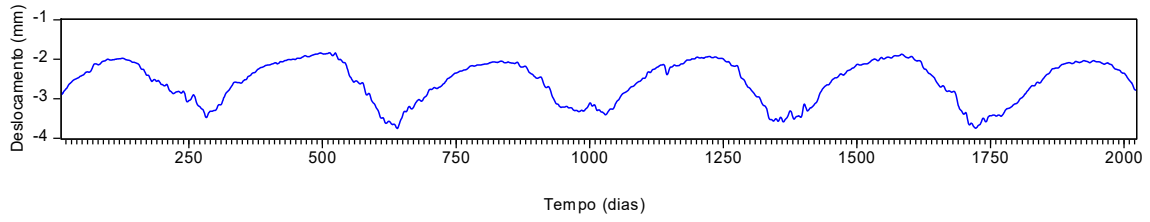
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 0.027	0.027	1.4287	0.232
		2 -0.009	-0.010	1.6085	0.447
		3 -0.023	-0.023	2.6849	0.443
		4 -0.019	-0.018	3.4378	0.487
		5 -0.010	-0.009	3.6385	0.603
		6 -0.000	-0.001	3.6386	0.725
		7 -0.027	-0.028	5.1151	0.646
		8 0.009	0.009	5.2669	0.729
		9 -0.015	-0.016	5.6974	0.770
		10 -0.013	-0.013	6.0247	0.813
		11 0.004	0.004	6.0564	0.870
		12 0.047	0.046	10.520	0.570
		13 0.090	0.087	26.893	0.013
		14 -0.012	-0.017	27.188	0.018
		15 0.012	0.017	27.478	0.025
		16 0.047	0.051	31.850	0.010
		17 0.018	0.019	32.473	0.013
		18 0.005	0.007	32.522	0.019
		19 0.024	0.029	33.708	0.020
		20 0.010	0.016	33.896	0.027
		21 0.045	0.047	37.971	0.013
		22 -0.012	-0.007	38.251	0.017
		23 -0.019	-0.012	38.974	0.020
		24 -0.011	-0.010	39.215	0.026
		25 -0.009	-0.014	39.396	0.034
		26 -0.003	-0.006	39.409	0.045
		27 0.029	0.030	41.111	0.040
		28 0.006	-0.002	41.173	0.052
		29 0.008	-0.003	41.300	0.065
		30 0.009	0.007	41.448	0.080
		31 -0.011	-0.014	41.694	0.095
		32 0.010	0.002	41.907	0.113
		33 0.019	0.011	42.662	0.121
		34 0.017	0.009	43.274	0.132
		35 0.008	0.009	43.405	0.156
		36 -0.001	0.001	43.406	0.185

Figura 46 – FAC e FACP dos resíduos quadráticos padronizados

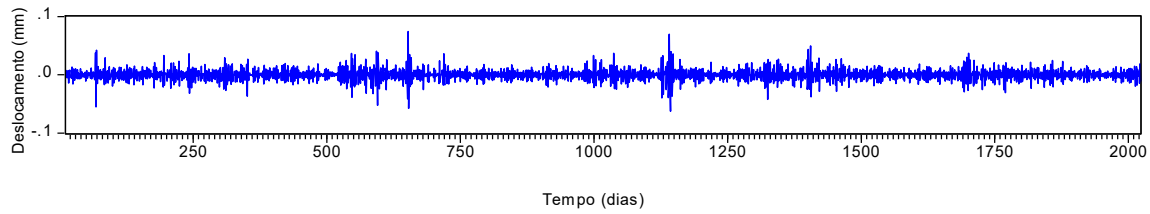
Vale ressaltar que a escolha do modelo ARIMA-GARCH, entre todos os modelos plausíveis identificados, foi determinada pela comparação do desempenho de previsão com o uso das estatísticas APE, MAPE, MAE e R^2 , em cada modelo candidato. O modelo selecionado foi aquele com menor APE, MAPE e MAE e com o maior R^2 .

4.2.3.2 Modelo WARIMAX-GARCH NEURAL

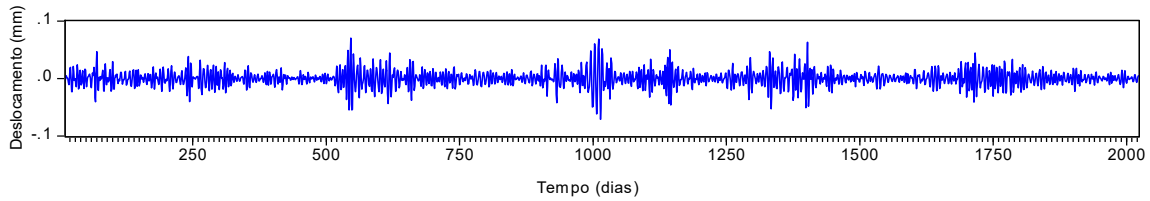
O método WARIMAX-GARCH Neural, conforme fluxograma da Figura 13, foi desenvolvido com base nos software MATLAB (versão 2013a) e EViews 8. Inicialmente, a decomposição *Wavelet* de nível $r = 2$ foi obtida a partir da amostra de treino da série de deslocamento, no sentido normal a montante-jusante, do bloco F13 da barragem principal da Usina Hidrelétrica de Itaipu. Os gráficos das componentes *Wavelet*, com base ortonormal *db40* (DAUBECHIES, 1988), podem ser vistos na Figura 47.



(a) Componente de aproximação de nível 2, $\tilde{y}_{A_2,t} (t = 1, \dots, 1993)$



(b) Componente de detalhe de nível 2, $\tilde{y}_{D_2,t} (t = 1, \dots, 1993)$



(c) Componente de detalhe de nível 3, $\tilde{y}_{D_3,t} (t = 1, \dots, 1993)$

Figura 47 – Componentes *Wavelet* da amostra de treino da série F13y

De acordo com a seção 2.1, para a decomposição *Wavelet* de nível $r = 2$, tem-se que $y_t = \tilde{y}_{A_2,t} + \tilde{y}_{D_2,t} + \tilde{y}_{D_3,t} + \varepsilon$, $(t = 1, \dots, 1993)$ onde ε é, estatisticamente, considerado ruído branco. Na segunda etapa do método WARIMAX-GARCH Neural, cada uma das componentes *Wavelet* $\tilde{y}_{A_2,t}$, $\tilde{y}_{D_2,t}$ e $\tilde{y}_{D_3,t}$ $(t = 1, \dots, 1993)$ foi modelada, individualmente, por RNA's e por modelos ARIMA-GARCH apropriados para cada série. Estes modelos foram utilizados para gerar as previsões fora da amostra para suas respectivas médias, denotadas por $\hat{y}_{A_2,t}$, $\hat{y}_{D_2,t}$ e $\hat{y}_{D_3,t}$ $(t = 1994, \dots, 2023)$. O horizonte de previsão escolhido foi $h = 30$, com 30 passos à frente, ou seja, foram previstos 30 dias de deslocamento do bloco F13 (no sentido normal ao do fluxo do rio), dado pela série F13y. Todos os parâmetros dos três modelos estimados são estatisticamente significativos, a 1%, e o diagnóstico residual confirma sua plausibilidade (Ver Apêndice IV D).

Numa terceira etapa, as três componentes *Wavelet* $\tilde{y}_{A_2,t,C}$, $\tilde{y}_{D_2,t,C}$ e $\tilde{y}_{D_3,t,C}$ (ou x_1, x_2 e x_3 , respectivamente) consistem de variáveis exógenas *Wavelet* geradas a partir da série de deslocamento F13y. Como mencionado anteriormente, elas são facilmente obtidas completando as componentes *Wavelet* com suas previsões fora da amostra, produzidas na segunda etapa.

Algebricamente, são representadas por:

$$\text{VII. } x_{1,t} = \tilde{y}_{A_2,t,C} := \left((\tilde{y}_{A_2,t})_{t=1}^{1993}; (\hat{y}_{A_2,t})_{t=1994}^{2023} \right);$$

$$\text{VIII. } x_{2,t} = \tilde{y}_{D_2,t,C} := \left((\tilde{y}_{D_2,t})_{t=1}^{1993}; (\hat{y}_{D_2,t})_{t=1994}^{2023} \right); e$$

$$\text{IX. } x_{3,t} = \tilde{y}_{D_3,t,C} := \left((\tilde{y}_{D_3,t})_{t=1}^{1993}; (\hat{y}_{D_3,t})_{t=1994}^{2023} \right);$$

Numa quarta etapa, foram usadas as variáveis exógenas *Wavelet* $x_{i,t}$ ($i = 1, 2, 3$) ($t = 1, \dots, 1993$) para o ajustamento do modelo WARIMAX-GARCH Neural utilizado para gerar as 30 previsões da série F13y (com 30 passos à frente). Algebricamente, a melhor formulação do modelo WARIMAX obtida foi:

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 + \phi_4 B^4 + \phi_5 B^5) y_t = (1 - \theta_1 B) e_t + \sum_{i=1}^2 \tilde{y}_{A_2,t-i,C} \varphi_{A_2,t-i} + \tilde{y}_{D_2,t-1,C} \varphi_{D_2,C,t-1} + \tilde{y}_{D_3,t-2,C} \varphi_{D_3,C,t-2} + \varphi \sigma_t^2.$$

Também, o melhor modelo GARCH obtido foi GARCH (1,1), com distribuição do erro generalizada (GED), com a formulação algébrica que segue:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

O método da Máxima Verossimilhança foi utilizado para obter as melhores estimativas dos parâmetros WARIMAX-GARCH.

No Apêndice IV E são exibidos os gráficos da FAC e FACP (para lags de 1 até 36) dos resíduos ordinários padronizados e resíduos quadráticos padronizados associados ao modelo WARIMAX-GARCH Neural. Nota-se que todos os valores da FAC e FACP pertencem ao intervalo de confiança com

99%, sugerindo que todos os valores são não significativos. A estatística de Ljung-Box (Q-Stat) sugere, também, que não há autodependência linear significativa, ao nível de 1% (dentro da amostra), para os resíduos ordinários e quadráticos padronizados provenientes do modelo WARIMAX-GARCH Neural. A estatística de Durbin-Watson, cujo valor obtido foi 1,986453, confirmou a ausência de autocorrelação de primeira ordem, dentro da amostra, para os resíduos padronizados.

Com base na Tabela 6, é possível comparar os resultados entre ARIMA-GARCH, RNA e WARIMAX-GARCH Neural dentro e fora da amostra de treino avaliando as estatísticas MAPE (*Mean Absolute Percent Error*) e MAE (*Mean Absolute Error*). Baseado nos dois tipos de erro, o modelo WARIMAX-GARCH Neural teve melhor desempenho que o modelo ARIMA-GARCH (5,0,3) x (2,1) e a RNA (ver configurações da RNA no Apêndice IV C), com base nas estatísticas MAPE, MAE e RMSE. Isso significa que o WARIMAX-GARCH Neural modelou melhor a dinâmica da série e também teve melhor poder de gerar valores (previsões) fora da amostra. A redução do erro obtido, fora da amostra, pelo método WARIMAX-GARCH Neural em relação ao modelo ARIMA-GARCH foi de até 83,8% e, comparado à RNA foi de até 84,1%, conforme Tabela 6.

MÉTODO	MAPE		MAE		RMSE	
	<i>Dentro da amostra</i>	<i>Fora da amostra</i>	<i>Dentro da amostra</i>	<i>Fora da amostra</i>	<i>Dentro da amostra</i>	<i>Fora da amostra</i>
ARIMA-GARCH	0,798%	9,627%	0,0207	0,2526	0,029	0,309
RNA	0,797%	10,012%	0,0207	0,2621	0,029	0,315
WARIMAX-GARCH Neural	0,457%	1,679%	0,0118	0,0435	0,016	0,050

Tabela 6 – Valores dos erros MAPE, MAE e RMSE, dentro e fora da amostra, para a série temporal F13Y

A Figura 48 mostra o gráfico do Erro Percentual Absoluto (APE) calculado fora da amostra para cada passo à frente ($t = 1994, \dots, 2023$) para os modelos ARIMA-GARCH e WARIMAX-GARCH Neural. Nota-se que o modelo WARIMAX-GARCH Neural possui melhores valores APE que o modelo ARIMA-

GARCH nos 30 valores fora da amostra, além de produzir previsões comparativamente melhores na medida em que o número de passos à frente foi aumentando. De fato, enquanto os erros de previsão do modelo ARIMA-GARCH mostram uma tendência positiva a partir do horizonte de previsão $h = 1$ (atingindo valores de até 20% para $h = 30$), o modelo WARIMAX-GARCH Neural mostra flutuações abaixo da linha de 3,2% com o aumento do horizonte de previsão. O Apêndice IV E mostra outros gráficos da evolução temporal fora da amostra e as respectivas previsões de ambas as abordagens para apoiar a conclusão de que, de fato, o modelo WARIMAX-GARCH Neural tem melhor poder de generalização.

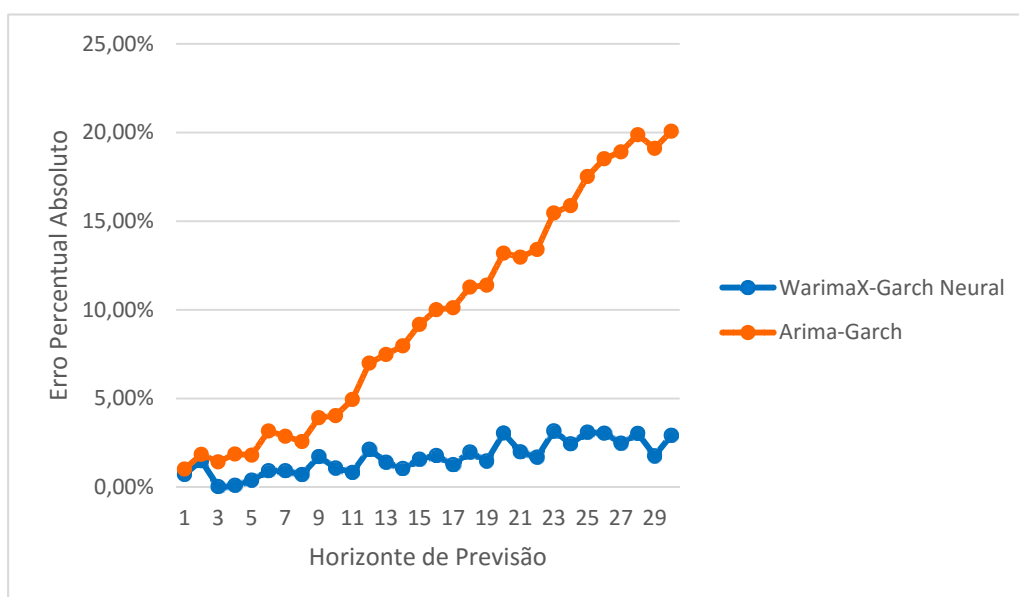


Figura 48 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH Neural e ARIMA-GARCH

São apresentados, a seguir, os gráficos dos deslocamentos reais e suas previsões, com intervalo de previsão de 99% para as previsões fora da amostra, obtidas por (a) Modelo ARIMA-GARCH (Figura 49) e (b) o Método WARIMAX-GARCH Neural (Figura 50). Nota-se, pela Figura 49, que as previsões obtidas pelo modelo ARIMA-GARCH foram maiores do que os valores observados em cada momento, no entanto, todos os valores

permaneceram dentro do intervalo de 99% de confiança para as previsões fora da amostra.

Por outro lado, na Figura 50, as previsões obtidas pelo método WARIMAX-GARCH Neural descreveram a trajetória do deslocamento mais próxima do deslocamento real e, também, todos os valores permaneceram dentro do intervalo de 99% de confiança para a previsão, cujos limites foram muito menores que os do modelo ARIMA-GARCH. Isto é, a variância da densidade preditiva do modelo WARIMAX-GARCH Neural foi menor. Este gráfico também mostra a dinâmica das previsões produzidas pelo método WARIMAX-GARCH Neural por meio das oscilações das previsões em torno dos valores reais de deslocamento.

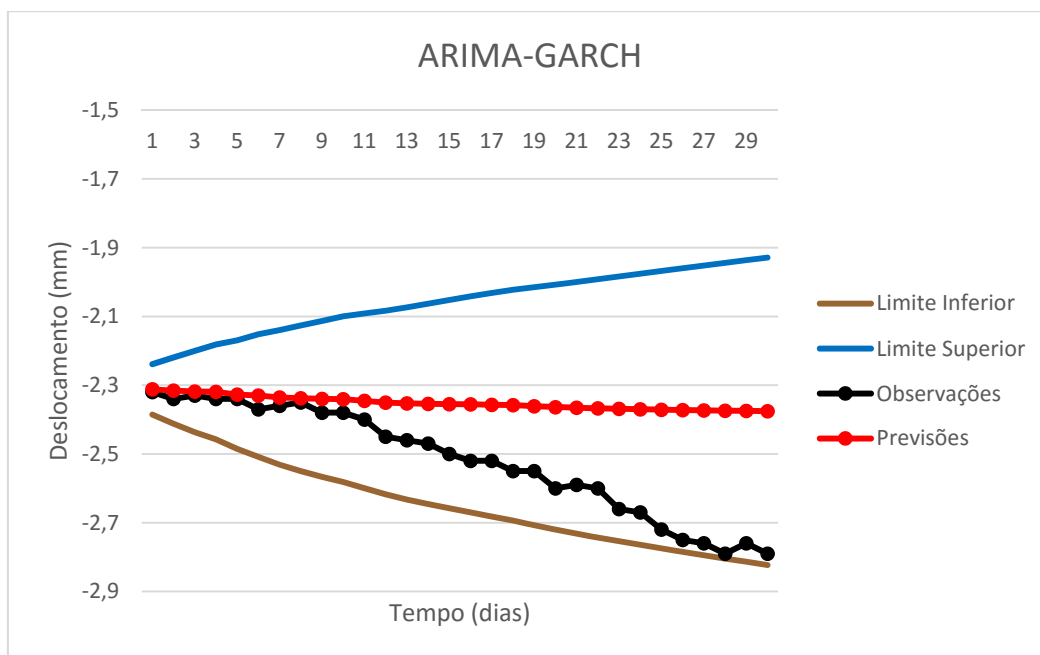


Figura 49 – Evolução temporal do intervalo de confiança para ARIMA-GARCH

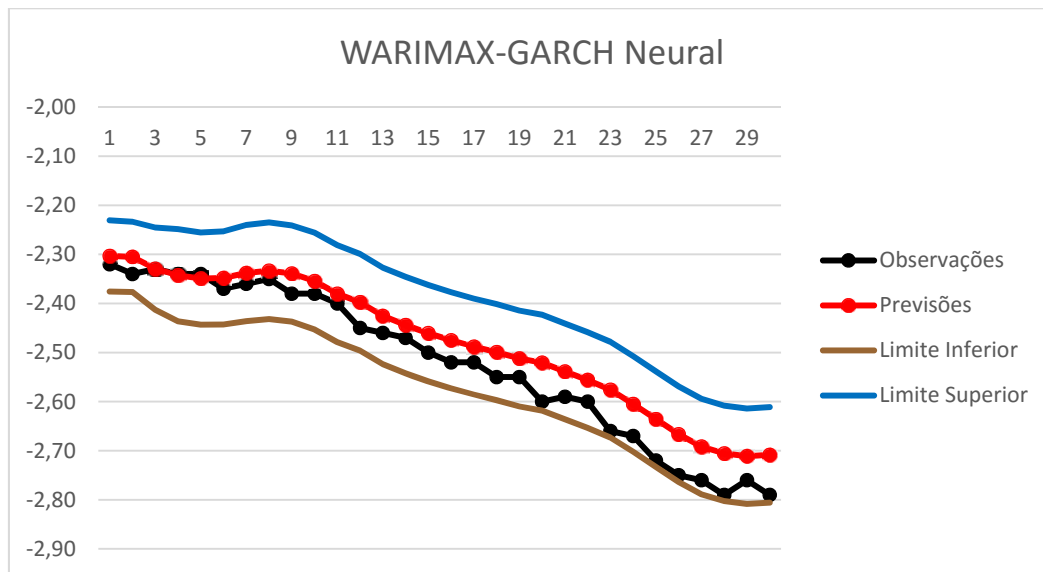


Figura 50 – Evolução Temporal do Intervalo de Confiança para WARIMAX-GARCH Neural

Com relação ao coeficiente de R^2 , que foi usado para mensurar a quantidade de variação nas séries temporais que é explicado pelo método/modelo estimado, o modelo ARIMA-GARCH teve $R^2 = 0,996960$ enquanto o modelo WARIMAX-GARCH Neural teve $R^2 = 0,760704$.

4.2.4 Método WARIMAX-GARCH NEURAL: Série F19x

Para a série F19x (ver descrição na Tabela 1, página 60) foram utilizadas 1993 observações (dias) para treinamento do modelo e 30 observações (dias) para teste, ou seja, horizonte de previsão $h = 30$, com 30 passos à frente (multipassos). Os valores aferidos pelo pêndulo automatizado, que compõe a série F19x representam o deslocamento do bloco F19, da barragem principal da Usina de Itaipu, no sentido do fluxo do rio. O gráfico desta série temporal é apresentado na Figura 51.

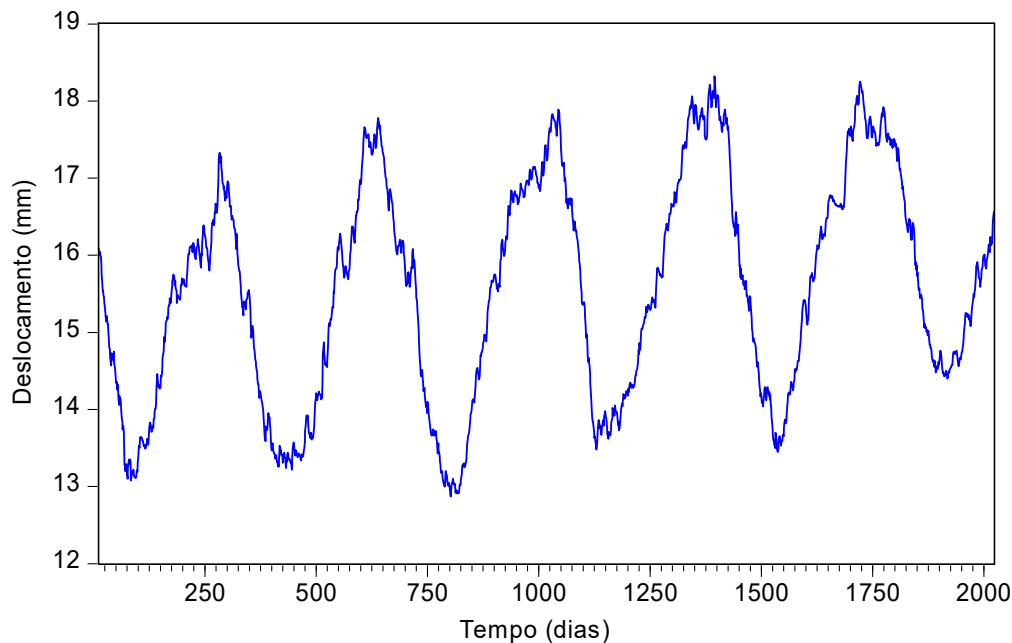


Figura 51 – Série temporal do deslocamento do bloco F19x

Na Figura 51 é possível notar que a série temporal exhibe oscilações de alta e baixa frequência nos dados do deslocamento. O gráfico da Figura 52, com os 90 últimos valores da série, evidencia mais explicitamente, as oscilações de alta frequência. Embora um modelo ARIMA sazonal seja apto a capturar as variações de baixa frequência da série temporal subjacente, em alguns casos, pode não ser capaz de modelar, de forma eficiente, as variações de alta frequência (representadas pelas componentes *Wavelet* de detalhes). Um modelo WARIMAX-GARCH Neural pode modelar ambas as variações no domínio do tempo da série temporal subjacente, através dos modelos ARIMA-GARCH integrados com a abordagem de decomposição *Wavelet*.

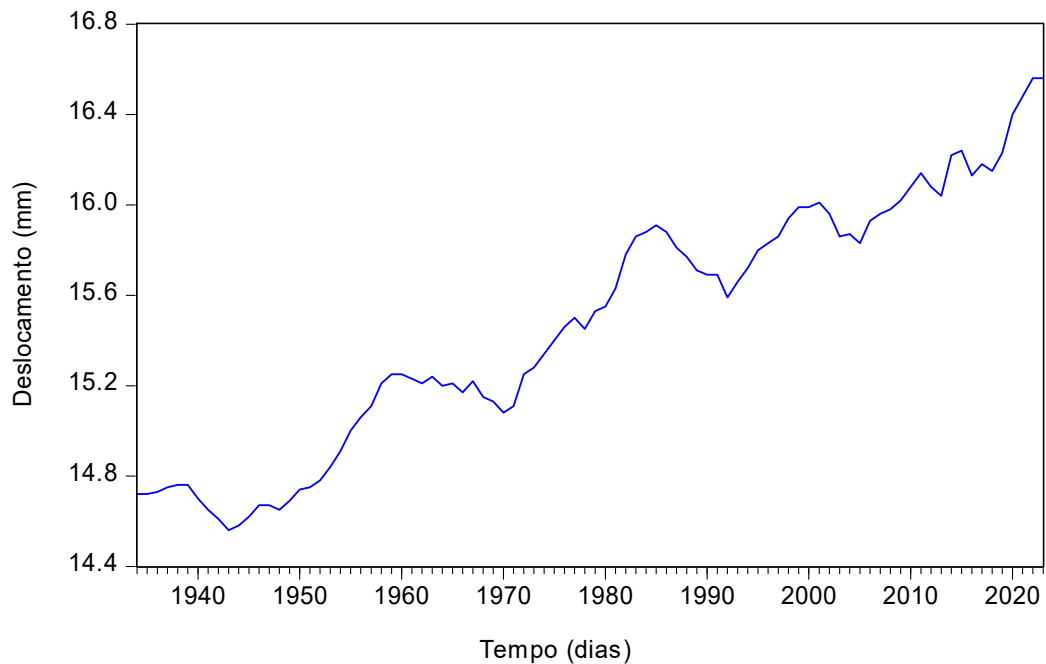


Figura 52 – Série temporal do Deslocamento do bloco F13x

O teste da raiz unitária Dickey-Fuller Aumentado (DFA) foi aplicado e não houve rejeição, ao nível de significância de 1%, da hipótese de não-estacionariedade dos dados de deslocamento. Porém, o teste DFA rejeitou fortemente a não estacionariedade da série temporal em sua segunda diferença (ver

Apêndice V A). As funções de autocorrelação simples e parciais, em relação a este último, bem como sua versão ao quadrado, apresentam valores significativos ao nível de 1%, o que sugere a escolha de um modelo ARIMA-GARCH. O modelo ARIMA-GARCH foi considerado plausível para o processo de modelagem, uma vez que não houve disponibilidade de variáveis exógenas convencionais apropriadas para serem utilizadas pelo modelo ARIMAX-GARCH para a série subjacente.

Os modelos plausíveis obtidos nesta seção geraram previsões pontuais e intervalares utilizando a metodologia 1-passo dentro da amostra e multipassos fora da amostra, tanto para o modelo ARIMA-GARCH quanto para o método WARIMAX-GARCH Neural. Foi utilizado o software EViews para realizar os diagnósticos residuais para a validação estatística com o teste DFA, e obter os testes de Ljung-Box e Durbin-Watson para autocorrelação de primeira ordem, bem como os gráficos das funções de autocorreção e autocorrelação parciais.

4.2.4.1 Modelo ARIMA-GARCH

Um modelo ARIMA (5,1,0) integrado com o modelo GARCH (1,1) (isto é, modelo ARIMA-GARCH (5,1,0) x (1,1)), erro com distribuição t de student, foi identificado para produzir o melhor ajuste para a amostra de treinamento y'_t ($t = 1, \dots, 1993$). Todos os parâmetros deste modelo foram estatisticamente significativos ao nível de 1%, e foram estimados por Máxima Verossimilhança.

Entre todos os modelos ARIMA-GARCH plausíveis, obtidos a partir da amostra de treinamento y'_t , foi o modelo ARIMA-GARCH (5,1,0) x (1,1) que produziu as melhores previsões para a média da série histórica subjacente, dentro da amostra de treinamento, em termos das estatísticas MAPE e MAE. Os resultados dos testes de Ljung-Box (Q-Stat) aplicados aos resíduos ordinários do modelo ARIMA-GARCH (5,1,0) x (1,1), conforme Figura 53 e Figura 54, sugerem que não há estrutura de autodependência linear significativa (ao nível de 1%) para os resíduos ordinários e quadráticos, até o $lag = 36$ (correspondendo a 36 dias). A Figura 54 confirma que não há autocorrelações significativas nos resíduos quadráticos padronizados (ao nível

de 1%) para os lags de 1 até 36 (Ver Apêndice V C). Além disso, o cálculo da estatística de Durbin-Watson, com valor 2,090102, sugere que não há autocorrelação de $lag = 1$ significativa nos resíduos ordinários.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	-0.024	-0.024	1.1737		
2	-0.008	-0.009	1.3072		
3	-0.022	-0.022	2.2443		
4	-0.006	-0.007	2.3059		
5	0.006	0.006	2.3825		
6	0.006	0.005	2.4450		
7	-0.022	-0.022	3.3773	0.066	
8	-0.009	-0.009	3.5238	0.172	
9	-0.017	-0.018	4.1192	0.249	
10	0.047	0.045	8.4855	0.075	
11	-0.021	-0.020	9.3973	0.094	
12	0.036	0.036	12.046	0.061	
13	-0.005	-0.002	12.094	0.097	
14	0.009	0.009	12.246	0.141	
15	0.008	0.009	12.389	0.192	
16	0.009	0.009	12.542	0.250	
17	0.005	0.008	12.597	0.320	
18	0.000	0.000	12.597	0.399	
19	0.054	0.057	18.408	0.143	
20	-0.002	-0.001	18.416	0.189	
21	0.053	0.058	23.964	0.066	
22	0.031	0.033	25.925	0.055	
23	0.008	0.015	26.060	0.073	
24	0.025	0.027	27.287	0.074	
25	0.005	0.008	27.330	0.097	
26	-0.013	-0.010	27.672	0.117	
27	0.034	0.034	29.937	0.093	
28	0.003	0.009	29.954	0.120	
29	0.028	0.026	31.517	0.111	
30	0.034	0.043	33.875	0.087	
31	0.024	0.021	35.005	0.088	
32	0.030	0.035	36.787	0.078	
33	-0.009	-0.008	36.939	0.096	
34	0.036	0.036	39.620	0.071	
35	0.011	0.013	39.870	0.086	
36	-0.012	-0.010	40.139	0.102	

Figura 53 – FAC e FACP dos resíduos ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	0.057	0.057	6.5604	0.010	
2	-0.012	-0.015	6.8300	0.033	
3	-0.020	-0.018	7.6129	0.055	
4	-0.031	-0.029	9.4929	0.050	
5	-0.027	-0.024	10.981	0.052	
6	-0.009	-0.007	11.127	0.085	
7	-0.016	-0.017	11.635	0.113	
8	-0.025	-0.025	12.839	0.118	
9	-0.027	-0.026	14.244	0.114	
10	-0.031	-0.031	16.181	0.095	
11	0.002	0.003	16.191	0.134	
12	0.001	-0.003	16.195	0.182	
13	-0.014	-0.018	16.590	0.219	
14	0.006	0.004	16.655	0.275	
15	0.059	0.056	23.651	0.071	
16	0.037	0.029	26.362	0.049	
17	0.007	0.002	26.465	0.066	
18	-0.006	-0.006	26.537	0.088	
19	0.010	0.015	26.750	0.111	
20	-0.029	-0.026	28.383	0.101	
21	-0.006	-0.000	28.446	0.128	
22	-0.026	-0.025	29.774	0.124	
23	0.005	0.010	29.820	0.155	
24	0.029	0.031	31.504	0.140	
25	0.009	0.009	31.668	0.168	
26	0.029	0.029	33.389	0.151	
27	-0.027	-0.031	34.873	0.142	
28	-0.007	0.000	34.963	0.171	
29	0.055	0.058	41.108	0.067	
30	0.031	0.020	42.986	0.059	
31	-0.010	-0.016	43.202	0.071	
32	0.011	0.014	43.452	0.085	
33	-0.009	-0.004	43.610	0.102	
34	0.016	0.024	44.143	0.114	
35	-0.003	-0.002	44.167	0.138	
36	-0.027	-0.023	45.683	0.129	

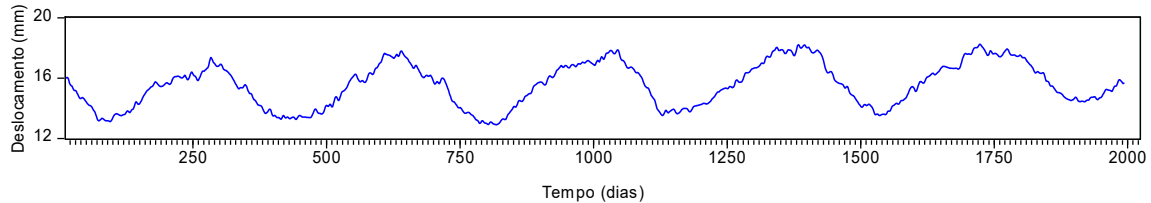
Figura 54 – FAC e FACP dos resíduos quadráticos padronizados

Vale ressaltar que a escolha do modelo ARIMA-GARCH, entre todos os modelos plausíveis identificados, foi determinada pela comparação do desempenho de previsão com o uso das estatísticas APE, MAPE, MAE e R^2 , em cada modelo candidato. O modelo selecionado foi aquele com menor APE, MAPE e MAE e com o maior R^2 .

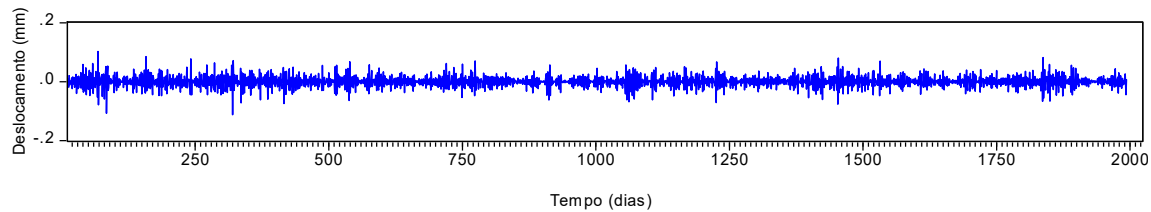
4.2.4.2 Modelo WARIMAX-GARCH NEURAL

O método WARIMAX-GARCH Neural, conforme fluxograma da Figura 13, foi desenvolvido com base nos software MATLAB (versão 2013a) e EViews 8. Inicialmente, a decomposição *Wavelet* de nível $r = 2$ foi obtida a partir da amostra de treino da série de deslocamento, no sentido montante-jusante, do bloco F19 da barragem principal da Usina Hidrelétrica de Itaipu. Os gráficos

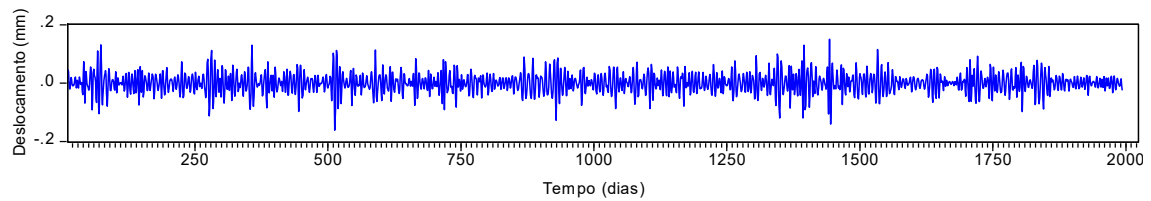
das componentes *Wavelet*, com base ortonormal *db40* (DAUBECHIES, 1988), podem ser vistos na Figura 55.



(a) Componente de aproximação de nível 2, $\tilde{y}_{A_2,t} (t = 1, \dots, 1993)$



(b) Componente de detalhe de nível 2, $\tilde{y}_{D_2,t} (t = 1, \dots, 1993)$



(c) Componente de detalhe de nível 3, $\tilde{y}_{D_3,t} (t = 1, \dots, 1993)$

Figura 55 – Componentes *Wavelet* da amostra de treino da série F19x

De acordo com a seção 2.1, para a decomposição *Wavelet* de nível $r = 2$, tem-se que $y_t = \tilde{y}_{A_2,t} + \tilde{y}_{D_2,t} + \tilde{y}_{D_3,t} + \varepsilon$, $(t = 1, \dots, 1993)$ onde ε é, estatisticamente, considerado ruído branco. Na segunda etapa do método WARIMAX-GARCH Neural, cada uma das componentes *Wavelet* $\tilde{y}_{A_2,t}$, $\tilde{y}_{D_2,t}$ e $\tilde{y}_{D_3,t}$ $(t = 1, \dots, 1993)$ foi modelada, individualmente, por RNA's e por modelos ARIMA-GARCH apropriados para cada série. Estes modelos foram utilizados para gerar as previsões fora da amostra para suas respectivas médias, denotadas por $\hat{y}_{A_2,t}$, $\hat{y}_{D_2,t}$ e $\hat{y}_{D_3,t}$ $(t = 1994, \dots, 2023)$. O horizonte de previsão escolhido foi $h = 30$, com 30 passos à frente, ou seja, foram previstos 30 dias de deslocamento do bloco F19 (no sentido normal ao do fluxo do rio), dado

pela série F19x. Todos os parâmetros dos três modelos estimados são estatisticamente significativos, a 1%, e o diagnóstico residual confirma sua plausibilidade (Ver Apêndice V D).

Numa terceira etapa, as três componentes *Wavelet* $\tilde{y}_{A_2,t,C}$, $\tilde{y}_{D_2,t,C}$ e $\tilde{y}_{D_3,t,C}$ (ou x_1, x_2 e x_3 , respectivamente) consistem de variáveis exógenas *Wavelet* geradas a partir da série de deslocamento F19x. Como mencionado anteriormente, elas são facilmente obtidas completando as componentes *Wavelet* com suas previsões fora da amostra, produzidas na segunda etapa.

Algebricamente, são representadas por:

$$\text{X. } x_{1,t} = \tilde{y}_{A_2,t,C} := \left((\tilde{y}_{A_2,t})_{t=1}^{1993}; (\hat{y}_{A_2,t})_{t=1994}^{2023} \right);$$

$$\text{XI. } x_{2,t} = \tilde{y}_{D_2,t,C} := \left((\tilde{y}_{D_2,t})_{t=1}^{1993}; (\hat{y}_{D_2,t})_{t=1994}^{2023} \right); \text{ e}$$

$$\text{XII. } x_{3,t} = \tilde{y}_{D_3,t,C} := \left((\tilde{y}_{D_3,t})_{t=1}^{1993}; (\hat{y}_{D_3,t})_{t=1994}^{2023} \right);$$

Numa quarta etapa, foram usadas as variáveis exógenas *Wavelet* $x_{i,t}$ ($i = 1, 2, 3$) ($t = 1, \dots, 1993$) para o ajustamento do modelo WARIMAX-GARCH Neural utilizado para gerar as 30 previsões da série F19x (com 30 passos à frente). Algebricamente, a melhor formulação do modelo WARIMAX obtida foi:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{14} B^{14}) y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_{14} B^{14}) e_t + \sum_{i=1}^2 \tilde{y}_{A_2,t-i,C} \varphi_{A_2,t-i} + \tilde{y}_{D_2,t-1,C} \varphi_{D_2,t-1} + \varphi \sigma_t^2.$$

Também, o melhor modelo GARCH obtido foi GARCH (1,1), com distribuição GED, com a formulação algébrica que segue:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

O método da Máxima Verossimilhança foi utilizado para obter as melhores estimativas dos parâmetros WARIMAX-GARCH.

No Apêndice V E são exibidos os gráficos da FAC e FACP (para lags de 1 até 36) dos resíduos ordinários padronizados e resíduos quadráticos

padronizados associados ao modelo WARIMAX-GARCH Neural. Nota-se que todos os valores da FAC e FACP pertencem ao intervalo de confiança com 99%, sugerindo que todos os valores são não significativos. A estatística de Ljung-Box (Q-Stat) sugere, também, que não há autodependência linear significativa, ao nível de 1% (dentro da amostra), para os resíduos ordinários e quadráticos padronizados provenientes do modelo WARIMAX-GARCH Neural. A estatística de Durbin-Watson, cujo valor obtido foi 2,030502, confirmou a ausência de autocorrelação de primeira ordem, dentro da amostra, para os resíduos padronizados.

Com base na Tabela 7, é possível comparar os resultados entre ARIMA-GARCH, RNA e WARIMAX-GARCH Neural dentro e fora da amostra de treino avaliando as estatísticas MAPE (Mean Absolute Percent *Error*) e MAE (Mean Absolute *Error*). Baseado nos dois tipos de erro, o modelo WARIMAX-GARCH Neural teve melhor desempenho que o modelo ARIMA-GARCH (5,1,0) x (1,1) e a RNA (ver configurações da RNA no Apêndice V C), com base nas estatísticas MAPE, MAE e RMSE. Isso significa que o WARIMAX-GARCH Neural modelou melhor a dinâmica da série e também teve melhor poder de gerar valores (previsões) fora da amostra. Conforme Tabela 7, o método WARIMAX-GARCH Neural reduziu em até 72,9% o erro fora da amostra quando comparado com o modelo ARIMA-GARCH e 73,4% quando comparado à RNA.

MÉTODO	MAPE		MAE		RMSE	
	<i>Dentro da amostra</i>	<i>Fora da amostra</i>	<i>Dentro da amostra</i>	<i>Fora da amostra</i>	<i>Dentro da amostra</i>	<i>Fora da amostra</i>
ARIMA-GARCH	0,313%	2,404%	0,0483	0,3890	0,064	0,444
RNA	0,317%	2,444%	0,0490	0,3951	0,064	0,437
WARIMAX-GARCH Neural	0,113%	0,651%	0,0204	0,1050	0,026	0,135

Tabela 7 – Valores dos erros MAPE, MAE e RMSE, dentro e fora da amostra, para a série temporal F19X

A Figura 56 mostra o gráfico do Erro Percentual Absoluto (APE) calculado fora da amostra para cada passo à frente ($t = 1994, \dots, 2023$) para os modelos ARIMA-GARCH e WARIMAX-GARCH Neural. Nota-se que o modelo

WARIMAX-GARCH Neural possui melhores valores APE que o modelo ARIMA-GARCH nos 30 valores fora da amostra, além de produzir previsões comparativamente melhores na medida em que o número de passos à frente vai aumentando. De fato, enquanto os erros de previsão do modelo ARIMA-GARCH mostram uma tendência positiva a partir do horizonte de previsão $h = 12$ (atingindo valores de até 5,32% para $h = 30$), o modelo WARIMAX-GARCH Neural mostra flutuações abaixo da linha de 2% com o aumento do horizonte de previsão. O Apêndice V E mostra outros gráficos da evolução temporal fora da amostra e as respectivas previsões de ambas as abordagens para apoiar a conclusão de que, de fato, o modelo WARIMAX-GARCH Neural tem melhor poder de generalização.

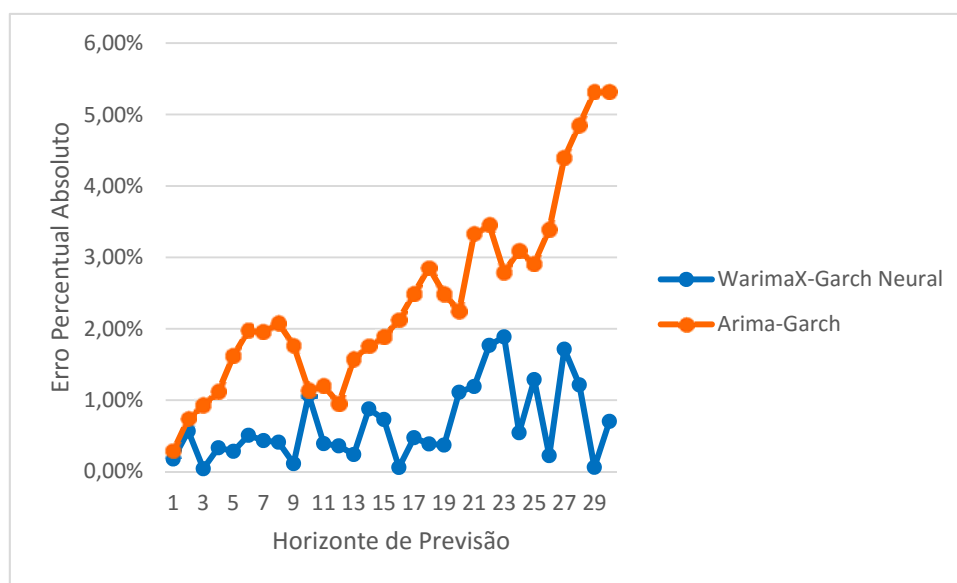


Figura 56 – Comparação temporal da evolução APE para os métodos WARIMAX-GARCH Neural e ARIMA-GARCH

São apresentados, a seguir, os gráficos dos deslocamentos reais e suas previsões, com intervalo de previsão de 99% para as previsões fora da amostra, obtidas por (a) Modelo ARIMA-GARCH (Figura 57) e (b) o Método WARIMAX-GARCH Neural (Figura 58). Nota-se, pela Figura 57, que as previsões obtidas pelo modelo ARIMA-GARCH foram maiores do que os valores observados em cada momento, no entanto, todos os valores

permaneceram dentro do intervalo de 99% de confiança para as previsões fora da amostra.

Por outro lado, na Figura 58, as previsões obtidas pelo método WARIMAX-GARCH Neural descreveram a trajetória do deslocamento mais próxima do deslocamento real e, também, todos os valores permaneceram dentro do intervalo de 99% de confiança para a previsão, cujos limites foram muito menores que os do modelo ARIMA-GARCH. Isto é, a variância da densidade preditiva do modelo WARIMAX-GARCH Neural foi menor. Este gráfico também mostra a dinâmica das previsões produzidas pelo método WARIMAX-GARCH Neural por meio das oscilações das previsões em torno dos valores reais de deslocamento.

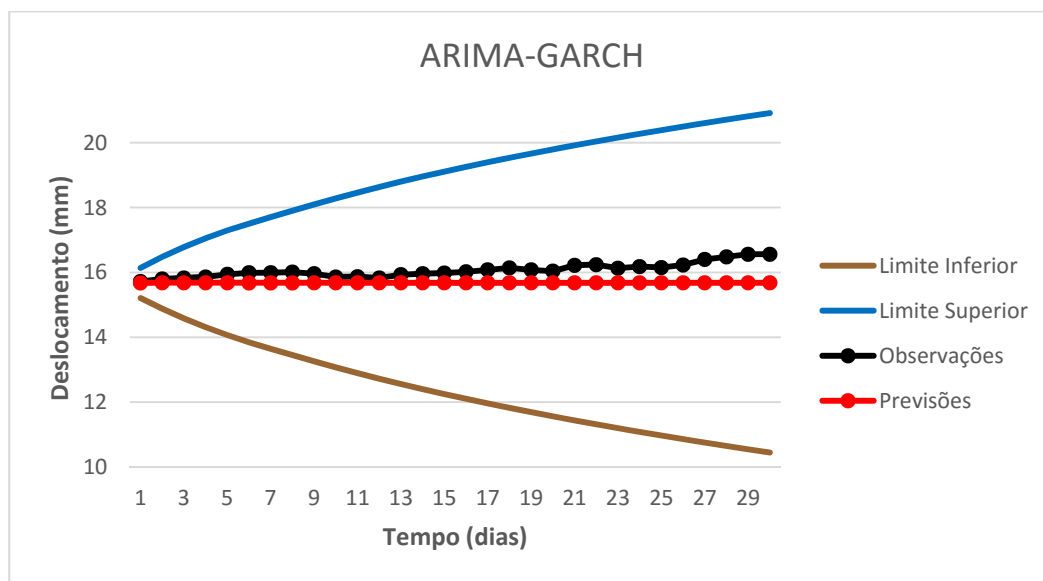


Figura 57 – Evolução temporal do intervalo de confiança para ARIMA-GARCH

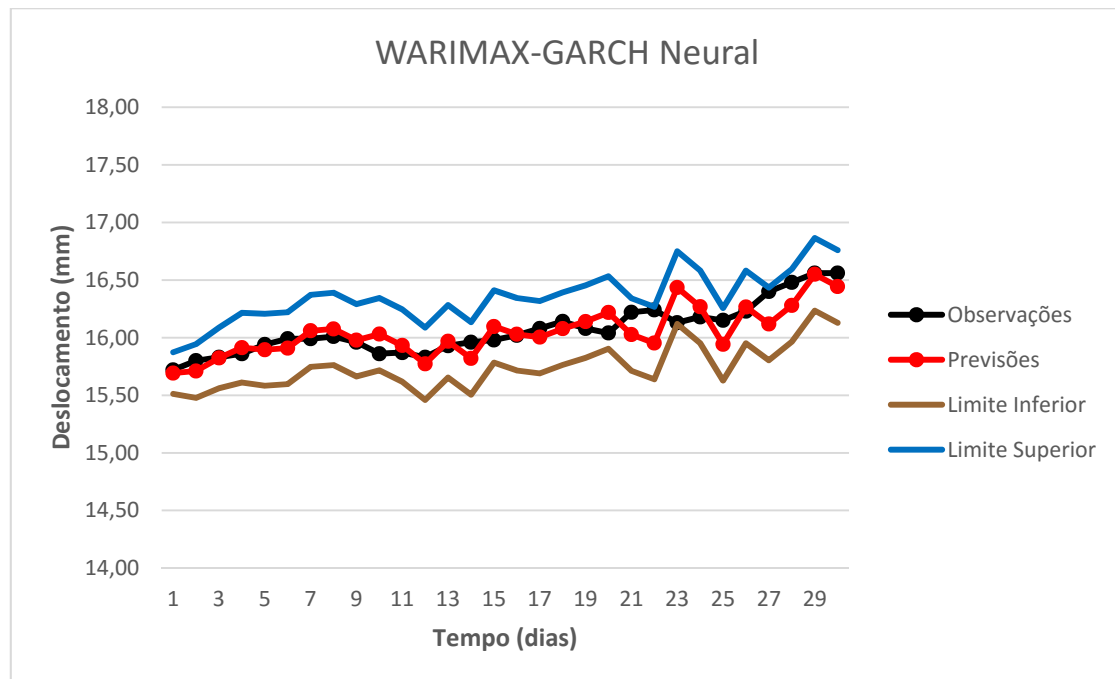


Figura 58 – Evolução Temporal do Intervalo de Confiança para WARIMAX-GARCH Neural

Com relação ao coeficiente R^2 , que foi usado para mensurar a quantidade de variação nas séries temporais que é explicado pelo método/modelo estimado, o modelo ARIMA-GARCH teve $R^2 = 0,173297$ enquanto o modelo WARIMAX-GARCH Neural teve $R^2 = 0,842793$.

5 CONCLUSÃO E CONSIDERAÇÕES FINAIS

Neste trabalho, o novo método de previsão causal chamado de método WARIMAX-GARCH Neural que incorpora variáveis exógenas *Wavelet* (obtidas a partir da decomposição *Wavelet* da série subjacente) mostrou, no estudo de caso, melhorias substanciais em relação aos modelos de previsão ARIMA-GARCH e às Redes Neurais Artificiais. As componentes *Wavelet* incorporadas têm boas propriedades estatísticas para serem utilizadas como variáveis exógenas pelo modelo WARIMAX-GARCH Neural. Por exemplo, as componentes de detalhe são sempre um processo estacionário de segunda ordem (usualmente exigido pelas variáveis exógenas que integram um modelo estatístico de regressão linear), estando sempre presente a variância condicional (volatilidade) – similarmente para um número da série histórica financeira (ver, por exemplo, Engle, 1982) – que permite que efeitos não lineares possam ser contabilizados no modelo final. Além disso, a componente *Wavelet* de aproximação sempre pode ser modelada por um ARIMA-GARCH sempre que a sua série de tempo original também seja capaz de ser modelada. Além disso, pode-se ver facilmente (e de forma intuitiva) que as componentes *Wavelet* apresentam sempre uma forte correlação com a variável resposta y_t ($t = 1, \dots, T$), uma vez que são obtidas a partir da própria série em estudo.

A utilização de variáveis exógenas híbridas, modeladas por ARIMA-GARCH e por RNA das componentes *Wavelet* da série original, foi responsável por capturar a estrutura de autodependência linear e não linear das componentes *Wavelet*, e tem se ajustado melhor no modelo WARIMAX-GARCH Neural, refletindo em ganhos preditivos quando comparado aos métodos clássicos ARIMA-GARCH e RNA, usados individualmente.

O método proposto foi aplicado a séries temporais diárias de deslocamentos horizontais de blocos da barragem da Usina Hidrelétrica de Itaipu, localizada na cidade de Foz do Iguaçu, Paraná, Brasil. A comparação dos resultados contra o modelo ARIMA-GARCH mostrou que o método WARIMAX-GARCH Neural não só produz performance de previsão

significativamente melhor, mas também produz previsões pontuais com intervalos de previsão associados de uma forma operacionalmente simples.

Vale salientar que o método WARIMAX-GARCH Neural não é um método exclusivo para previsões de séries temporais de pêndulos, e pode ser utilizado em séries temporais com estrutura de autodependência linear e não linear, e também em séries temporais que exibem volatilidade.

6 REFERÊNCIAS

- ABELÉM, A. **Redes Neurais Artificiais na Previsão de Séries Temporais**, 1994. Departamento de Engenharia Elétrica. Dissertação de Mestrado. Pontifícia Universidade Católica do Rio de Janeiro.
- AN, X.; JIANG, D.; LIU, C.; ZHAO, M. Wind farm power prediction based on wavelet decomposition and chaotic time series. **Expert Systems with Applications**, v. 38, n. 9, p. 11280–11285, 2011. Elsevier Ltd.
- ANEEL, A. N. D. E. E. **Atlas de energia elétrica do Brasil**. Agência Nacional de Energia, Brasília, 2008.
- BATES, J. M.; GRANGER, C. W. J. The Combination of Forecasts. **Journal of the Operational Research Society**, v. 20, n. 4, p. 451–468, 1969. Nature Publishing Group.
- BENAOUDA, D.; MURTAGH, F.; STARCK, J. L.; RENAUD, O. Wavelet-based nonlinear multiscale decomposition model for electricity load forecasting. **Neurocomputing**, v. 70, p. 139–154, 2006.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, v. 31, p. 307–327, 1986.
- BOX, G. E.; JENKINS, G. M. **Time Series Analysis: Forecasting and Control**. Oakland, Califórnia: Holden-Day, 1970.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. 4th ed. ed. Hoboken, NJ: Wiley, 2008.
- BOX, G. E. P.; TIAO, G. C. Intervention Analysis with Applications to Economic and Environmental problems. **Journal of American Statistical Association**, v. 70, n. 349, p. 70–79, 1975.
- BRASIL, M. DE S. E I. DE B. **Manual de Segurança e Inspeção de Barragens**. 2002.
- BUNN, D. W. Combining forecasts. **European Journal of Operational Research**, v. 33, n. 3, p. 223–229, 1988.
- BYUN, S. J.; CHO, H. Forecasting carbon futures volatility using GARCH models with energy volatilities. **Energy Economics**, v. 40, p. 207–221, 2013. Elsevier B.V.
- CAMPOS, L. C. D. **Modelo Estocástico Periódico baseado em Redes Neurais**, 2010. Tese de Doutorado. Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio).
- CAVALERI, R.; RIBEIRO, E. P. Combinação de Previsões de Volatilidade : Um Estudo. **Revista Economia**, , n. 1984, 2011.
- CHANG, F.-J.; CHEN, P.-A.; LU, Y.-R.; HUANG, E.; CHANG, K.-Y. Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban

- flood control. **Journal of Hydrology**, v. 517, p. 836–846, 2014. Elsevier B.V.
- CHARYTONIUK, W.; NIEBRZYDOWSKI, J. Confidence interval construction for load forecast. **Electric Power Systems Research**, v. 48, p. 97–103, 1998.
- CHLAPINKSI, G.; RÓZANSKI, R. Prediction intervals for time series models with trend via sieve bootstrap. **Journal of Statistical Planning and Inference**, v. 143, p. 221–236, 2013.
- CLEMEN, R. T. Combining forecasts: A review and annotated bibliography. **International Journal of Forecasting**, v. 5, p. 559–583, 1989.
- CLEMENTS, M. P.; KIM, J. H. Bootstrap prediction intervals for autoregressive time series. **Computational Statistics and Data Analysis**, v. 51, p. 3580–3594, 2007.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of Control, Signals, and Systems**, v. 2, n. 4, p. 303–314, 1989.
- DAUBECHIES, I. Orthonormal bases of compactly supported wavelets. **Communications on Pure and Applied Mathematics**, v. 41, n. 7, p. 909–996, 1988.
- DAUBECHIES, I. The wavelet transform, time-frequency localization and signal analysis. **IEEE Transactions on Information Theory**, v. 36, n. 5, p. 961–1005, 1990.
- DONOHU, D. L.; JOHNSTONE, J. M. Ideal spatial adaptation by wavelet shrinkage. **Biometrika**, v. 81, n. 3, p. 425–455, 1994.
- ENGLE, R. F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. **Econometrica**, v. vol. 50, n. No. 4, p. 987–1007, 1982.
- FARIA, A. E.;; MUBWANDARIKWA, E. Multimodality on the Geometric Combination of Bayesian Forecasting Models. **International Journal of Statistics and Management System**, v. 3, p. 1–25, 2008.
- FERNANDES, C. **Notas de aula do curso de Séries Financeiras**. Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro. 2006, 2006.
- FIORDALISO, A. A nonlinear forecasts combination method based on Takagi–Sugeno fuzzy systems. **International Journal of Forecasting**, v. 14, n. 3, p. 367–379, 1998.
- GABRIEL, A. S. Evaluating the Forecasting Performance of GARCH Models. Evidence from Romania. **Procedia - Social and Behavioral Sciences**, v. 62, p. 1006–1010, 2012.
- DE GOOIJER, J. G.; HYNDMAN, R. J. 25 Years of Time Series Forecasting. **International Journal of Forecasting**, v. 22, n. 44, p. 443–473, 2006.
- GRANGER, C. W. J. Invited review combining forecasts—twenty years later. **Journal of Forecasting**, v. 8, n. 3, p. 167–173, 1989.

GRANGER, C. W. J.; RAMANATHAN, R. Improved Methods of Forecasting. **Journal of Forecasting**, v. 3, p. 197–204, 1984.

GRAPS, A. An Introduction to Wavelets. **IEEE Computational Science and Engineering**, v. 2, n. 2, p. 50–61, 1995.

GUPTA, S.; WILTON, P. C. Combination of Forecasts: An Extension. **Management Science**, v. 33, n. 3, p. 356–372, 1987. INFORMS.

HAAR, A. Zur Theorie der orthogonalen Funktionensysteme. **Mathematische Annalen**, v. 69, n. 3, p. 331–371, 1910.

HAMILTON, J. **Time Series Analysis**. Princeton University Press, 1994.

HAVEN, E.; LIU, X.; SHEN, L. De-noising option prices with wavelet method. **European Journal of Operational Research**, v. 222, p. 104–112, 2012.

HAYKIN, S. **Redes Neurais: Princípios e Prática**. 2º ed. Porto Alegre, Bookman, 2001.

HERRERA, F.; LOZANO, M. V. J. L. Artificial Intelligence Review. **Tackling Real-coded Genetic Algorithms: Operators and Tools for Behavioral Analysis**, journal, p. 265–319, 1998.

HOLLAUER, G.; ISSLER, J. V.; NOTINI, H. H. Prevendo o crescimento da produção industrial usando um número limitado de combinações de previsões. **Economia Aplicada**, v. 12, n. 2, p. 177–198, 2008. Economia Aplicada.

HSU, C.-C.; CHEN, C.-Y. Regional load forecasting in Taiwan—applications of artificial neural networks. **Energy Conversion and Management**, v. 44, p. 1941–1949, 2003.

ITAIPU. **Usina Hidrelétrica de Itaipu: Aspectos de Engenharia**. Foz do Iguaçu: TAB Marketing Editorial, 2009.

ITAIPU. Barragem Itaipu. Disponível em: <<https://www.itaipu.gov.br/energia/barragem>>. Acesso em: 20/2/2015a.

ITAIPU. Instrumentação. Disponível em: <<https://www.itaipu.gov.br/energia/instrumentacao>>. Acesso em: 20/2/2015b.

ITAIPU. Segurança de Barragens. Disponível em: <http://jie.itaipu.gov.br/print_node.php?secao=turbinadas1&nid=11134>. Acesso em: 20/2/2015c.

JAMMAZI, R.; ALOUI, C. Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. **Energy Economics**, v. 34, n. 3, p. 828–841, 2012. Elsevier B.V.

KARTHIKEYAN, L.; KUMAR, D. N. Predictability of nonstationary time series using wavelet and EMD based ARMA models. **Journal of Hydrology**, v. 502, p. 103–119, 2013.

KISI, O. Wavelet regression model for short-term streamflow forecasting.

Journal of Hydrology, v. 389, n. 3-4, p. 344–353, 2010.

KISI, O.; CIMEN, M. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. **Journal of Hydrology**, v. 399, n. 1-2, p. 132–140, 2011.

KUBRUSLY, C. S. **Elements of Operator Theory**. Birkhäuser, Boston, 2001.

KUBRUSLY, C. S.; LEVAN, N. Abstract Wavelets Generated by Hilbert Space Shift Operators. **Adv. Math. Sci. Appl.**, v. 14, p. 643–660, 2006.

KURT, A.; OKTAY, A. B. Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. **Expert Systems with Applications**, v. 37, p. 7986–7992, 2010.

LEVAN, N.; KUBRUSLY, C. S. A wavelet “time-shift-detail” decomposition. **Mathematics and Computers in Simulation**, v. 63, p. 73–78, 2003.

LI, W.; WANG, W.; WANG, X.; et al. A dynamic relearning neural network model for time series analysis of online marine data. **Computers & Geosciences**, v. 73, p. 99–107, 2014. Elsevier.

LIMA, E. L. **Análise Real. Funções de uma Variável**. Volume 1, 8º Edição, Rio de Janeiro: IMPA, 2006.

LIMA, P. C. Wavelets: uma introdução. **Matemática Universitária**, v. 33, p. 13–44, 2002.

LIMA, P. C. DE. Wavelets: uma introdução. , p. 13–44, 2002.

LIU, H.; SHI, J. Applying ARMA-GARCH approaches to forecasting short-term electricity prices. **Energy Economics**, v. 37, p. 152–166, 2013. Elsevier B.V.

LIU, L. M. **Time Series Analysis and Forecasting**. Scientific Computing Associations, 2006.

LUTKEPOHL, H. **New Introduction to Multiple Time Series Analysis**. Springer, 2006.

MALLAT, S. **A Wavelet Tour of Signal Processing: The Sparse Way**. Third Edit ed. 2008.

MALLAT, S. G. Theory for multiresolution signal decomposition: the wavelet representation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 11, n. 7, p. 674–693, 1989.

MARQUARDT, D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. **SIAM Journal on Applied Mathematics**, v. 11, n. 2, p. 431–441, 1963.

MARTINS, V. L. M. **Comparação de combinação de previsões correlacionadas e não correlacionadas com as suas previsões individuais: um estudo com séries industriais**, 2011. Porto Alegre: Universidade Federal do Rio Grande do Sul.

MATTAR, M. A.; ALAZBA, A. A.; ZIN EL-ABEDIN, T. K. Forecasting furrow irrigation infiltration using artificial neural networks. **Agricultural Water Management**, v. 148, p. 63–71, 2015. Elsevier B.V.

MILLAR, R. B. **Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB**. Wiley, 2011.

MISIOREK, A.; TRUECK, S.; WERON, R. Point and Interval Forecasting of Spot Electricity Prices: Linear vs. Non-Linear Time Series Models. **Studies in Nonlinear Dynamics & Econometrics**, 20. jun. 2006. The Berkeley Electronic Press.

MORETTIN, P. A.; TOLOI, L. M. C. **Análise Séries Temporais**. 2ª Ed. ABE Projeto Fisher, 2006.

NALLEY, D.; ADAMOWSKI, J.; KHALIL, B. Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008). **Journal of Hydrology**, v. 475, p. 204–228, 2012.

NEWBOLD, P.; GRANGER, J. C. W. Experience with Forecasting Univariate Time Series and the Combination of Forecasts. **Journal of the Royal Statistical Society. Series A (General)**, v. 137, p. 131–165, 1974.

O'CONNOR, M.; LAWRENCE, M. An examination of the accuracy of judgemental confidence intervals in time series forecasting. **Journal of Forecasting**, v. 8, n. 2, p. 141–155, 1989.

PALIT, A. K.; POPOVIC, D. **Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications**. Springer-Verlag, London., 2005.

PANKRATZ, A. **Forecasting with Dynamic Regression Models**. Wiley-Interscience., 1991.

RAGSDALE, C. **Spreadsheet Modeling & Decision Analysis: A Practical Introduction to Management Science**. Fourth edition, South-Western, 2004.

RAMSEY, J. B.; GALLEGATI, M.; GALLEGATI, M.; SEMMLER, W. Instrumental variables and wavelet decompositions. **Economic Modelling**, v. 27, n. 6, p. 1498–1513, 2010.

REBOREDO, J. C.; RIVERA-CASTRO, M. A. A wavelet decomposition approach to crude oil price and exchange rate dependence. **Economic Modelling**, v. 32, p. 42–57, 2013. Elsevier B.V.

REZENDE, S. O. **Sistemas inteligentes: Fundamentos e Aplicações**. Editora Manole Ltda, 2003.

RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. New Jersey: Prentice-Hall Inc, 1995.

SEO, Y.; KIM, S.; KISI, O.; SINGH, V. P. Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. **Journal of**

Hydrology, v. 520, p. 224–243, 2015. Elsevier B.V.

SILVEIRA, J. F. A. REVISÃO CRÍTICA SOBRE A INSTRUMENTAÇÃO DE BARRAGENS NO BRASIL. **Comitê Brasileiro De Barragens**, v. XXIX SEMIN, 2013.

SLUTZKY, E. The Summation of Random Causes as the Source of Cyclic Processes. **Econometrica**, v. 5, n. 2, p. 105, 1937.

SOLTANI, S.; SOLTANI, S. On the use of the wavelet decomposition for time series prediction. , v. 48, p. 267–277, 2002.

TAFNER, M. A. Redes Neurais Artificiais: Introdução e Princípios de Neurocomputação. , 1996. Blumenau: EKO.

TEIXEIRA JÚNIOR, L. A. **Combinação SSA-Wavelet de Métodos Preditivos com Ajuste Numérico MINIMAX na Geração de Previsões e de Cenários**, 2013.

TEIXEIRA JÚNIOR, L. A.; PESSANHA, J. F. M.; MENEZES, M. L. DE; CASSIANO, K. M.; SOUZA, R. C. Redes neurais artificiais wavelet causal na previsão da radiação solar direta. **Simpósio Brasileiro de Pesquisa Operacional**, 2012.

TEIXEIRA JÚNIOR, L. A.; SOUZA, R. M.; MENEZES, L. M.; et al. Artificial Neural Network and Wavelet decomposition in the Forecast of Global Horizontal Solar Radiation. **Sobrapo**, v. 35, n. 1, p. 1–16, 2015.

TIWARI, M. K.; CHATTERJEE, C. Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN (WBANN) hybrid approach. **Journal of Hydrology**, v. 394, n. 3-4, p. 458–470, 2010.

VANDERPLAATS, G. N. **Numerical Optimization Techniques for Engineering Design**. Third edition, Vanderplaats Research & Development, Inc, Colorado Springs., 1999.

WALLIS, K. F. Combining forecasts – forty years later. **Applied Financial Economics**, v. 21, n. 1-2, p. 33–41, 2011. Routledge.

WERNER, L.; RIBEIRO, J. L. D. Modelo composto para prever demanda através da integração de previsões. , 2006.

WINKLER, R. L.; MAKRIDAKIS, S. The Combination of Forecasts. **Journal of the Royal Statistical Society. Series A (General)**, v. 146, p. 150–157, 1983.

ZOU, H.; YANG, Y. Combining time series models for forecasting. **International Journal of Forecasting**, v. 20, p. 69–84, 2004.

ZSOLT, L. **Redes Neurais Artificiais: Fundamentos e Aplicações**. 4º Ed, Livraria de Física, 2006.

7 APÊNDICES

Apêndice I A: Teste Dikey-Fuller Aumentado (série temporal F05x).

Null Hypothesis: DISPLACEMENT has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.093416	0.2475
Test critical values: 1% level	-3.432770	
5% level	-2.862495	
10% level	-2.567323	

Null Hypothesis: D(DISPLACEMENT) has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-28.37689	0.0000
Test critical values: 1% level	-3.432770	
5% level	-2.862495	
10% level	-2.567323	

Null Hypothesis: D(log(DISPLACEMENT),2) has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-21.01401	0.0000
Test critical values: 1% level	-3.432788	
5% level	-2.862503	
10% level	-2.567328	

Apêndice I B: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (série temporal F05x).

O modelo ARIMA (3,2,4), com a transformação logaritmo, é dado por:

$$\nabla^2 \left(\ln(\hat{f}(t)) \right) = [\tilde{f}(t-1) \times 0.370311990136 - \tilde{f}(t-3) \times 0.0429434724332] + [-\hat{e}(t-1) \times 0.928193311211 - \hat{e}(t-4) \times 0.066815164786]$$

onde ∇^2 denota o operador diferença de segunda ordem não sazonal. Por sua vez, o modelo GARCH (2,1) é dado por:

$$\hat{\sigma}_t^2 = 6.39028337782 \times 10^{-7} + (\hat{e}(t-1))^2 \times 0.298783476992 - (\hat{e}(t-2))^2 \times 0.271843956286 + \hat{\sigma}_{A_2t-1}^2 \times 0.944774588475.$$

Dependent Variable: D(D(LOG(F05X)))

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	0.049165	0.014277	3.443722	0.0006
AR(1)	0.370312	0.018638	19.86866	0.0000
AR(3)	-0.042943	0.015477	-2.774594	0.0055
MA(1)	-0.928193	0.003914	-237.1308	0.0000
MA(4)	-0.066815	0.004387	-15.23099	0.0000
Variance Equation				
C	6.39E-07	1.24E-07	5.143551	0.0000
RESID(-1)^2	0.298783	0.039504	7.563335	0.0000
RESID(-2)^2	-0.271844	0.038146	-7.126483	0.0000
GARCH(-1)	0.944775	0.009957	94.88592	0.0000
GED PARAMETER	1.271585	0.012687	100.2276	0.0000
R-squared	0.361378	Mean dependent var	3.44E-07	
Adjusted R-squared	0.360354	S.D. dependent var	0.009049	
S.E. of regression	0.007237	Akaike info criterion	-7.655592	
Sum squared resid	0.130685	Schwarz criterion	-7.632296	
Log likelihood	9579.490	Hannan-Quinn criter.	-7.647135	
Durbin-Watson stat	2.492367			

Apêndice I C: Modelo WARIMAX-GARCH e seus resultados estatísticos para validação (Série temporal F05x).

(1) Modelo ARIMA-GARCH para a componente de aproximação $\left(\tilde{f}_{A_2}(t)\right)_{t=1}^{2506}$,

onde o modelo ARIMA é dado por:

$$\ln\left(\hat{f}_{A_2}(t)\right) = [\tilde{f}_{A_2}(t-1) \times 0.125897580821 + \tilde{f}_{A_2}(t-2) \times 0.125770512376 + \tilde{f}_{A_2}(t-3) \times 0.125709093625 + \tilde{f}_{A_2}(t-4) \times 0.125628835092 + \tilde{f}_{A_2}(t-8) \times 0.125271106733 + \tilde{f}_{A_2}(t-13) \times 0.125109696855 + \tilde{f}_{A_2}(t-16) \times 0.125154812097 + \tilde{f}_{A_2}(t-20) \times 0.125095166896] + [\hat{e}_{A_2}(t -$$

$$1) \times 0.411740575818 + \hat{e}_{A_2}(t-2) \times 0.378547737351 + \hat{e}_{A_2}(t-4) \times 0.364247927235 + \hat{e}_{A_2}(t-8) \times 0.338461318146 + \hat{e}_{A_2}(t-12) \times 0.343107563218 + \hat{e}_{A_2}(t-20) \times 0.265105576746 + 0 \times \hat{\sigma}_{A_2,t}^2 \Big]$$

onde $\ln(.)$ é a transformação logaritmo e o modelo GARCH (1,1), com distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{A_2,t}^2 = 0.0824645040406 - 1.92575343596 \times (\hat{e}_{A_2}(t-1))^2 - 0.995804205137 \hat{\sigma}_{A_2,t-1}^2.$$

Dependent Variable: LOG(A2_F05X)

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(3)	0.125744	0.004220	29.79552	0.0000
AR(2)	0.125810	0.003505	35.89790	0.0000
AR(1)	0.125942	0.005031	25.03381	0.0000
AR(4)	0.125649	0.004513	27.84414	0.0000
AR(8)	0.125279	0.005309	23.59815	0.0000
AR(13)	0.125100	0.003076	40.67111	0.0000
AR(16)	0.125155	0.004425	28.28257	0.0000
AR(20)	0.125071	0.002625	47.65432	0.0000
MA(1)	0.423769	0.002983	142.0814	0.0000
MA(2)	0.386896	0.002221	174.2364	0.0000
MA(4)	0.363700	0.003183	114.2701	0.0000
MA(8)	0.336198	0.002924	114.9686	0.0000
MA(20)	0.255157	0.002856	89.35024	0.0000
MA(12)	0.341672	0.001904	179.4763	0.0000
Variance Equation				
C	0.017180	0.005247	3.274475	0.0011
RESID(-1)^2	-0.715356	0.271351	-2.636271	0.0084
GARCH(-1)	-0.882933	0.181478	-4.865245	0.0000
GED PARAMETER	0.254503	0.012735	19.98441	0.0000
R-squared	0.991857	Mean dependent var	2.424472	
Adjusted R-squared	0.991814	S.D. dependent var	0.121061	
S.E. of regression	0.010953	Akaike info criterion	-5.568893	
Sum squared resid	0.296440	Schwarz criterion	-5.526751	
Log likelihood	6937.350	Hannan-Quinn criter.	-5.553589	
Durbin-Watson stat	1.991000			

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.002	0.002	0.0072		
2	0.001	0.001	0.0116		
3	0.000	0.000	0.0117		
4	0.004	0.004	0.0572		
5	-0.005	-0.005	0.1094		
6	-0.000	-0.000	0.1094		
7	0.001	0.001	0.1144		
8	0.003	0.003	0.1363		
9	-0.008	-0.008	0.3097		
10	-0.005	-0.005	0.3774		
11	-0.003	-0.003	0.4000		
12	-0.015	-0.015	0.9741		
13	-0.001	-0.001	0.9757	0.323	
14	-0.003	-0.003	0.9966	0.608	
15	-0.006	-0.006	1.0954	0.778	
16	-0.034	-0.034	4.0621	0.398	
17	-0.024	-0.024	5.5478	0.353	
18	-0.008	-0.008	5.7015	0.457	
19	-0.003	-0.003	5.7244	0.572	
20	0.033	0.034	8.5748	0.379	
21	0.043	0.043	13.291	0.150	
22	0.005	0.005	13.360	0.204	
23	-0.002	-0.002	13.371	0.270	
24	-0.013	-0.013	13.791	0.314	
25	0.018	0.018	14.632	0.331	
26	0.021	0.020	15.715	0.331	
27	0.010	0.009	15.953	0.385	
28	0.068	0.066	27.668	0.035	
29	0.027	0.027	29.596	0.029	
30	0.010	0.010	29.838	0.039	
31	0.011	0.011	30.131	0.050	
32	0.041	0.041	34.518	0.023	
33	0.015	0.015	35.095	0.028	
34	0.014	0.014	35.618	0.033	
35	0.008	0.009	35.774	0.043	
36	-0.020	-0.018	36.832	0.045	

FAC e FACP dos resíduos
ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	0.003	0.003	0.0293	0.864	
2	0.005	0.005	0.0958	0.953	
3	-0.000	-0.000	0.0960	0.992	
4	0.006	0.006	0.1949	0.996	
5	0.005	0.005	0.2550	0.998	
6	0.000	-0.000	0.2550	1.000	
7	-0.000	-0.000	0.2555	1.000	
8	-0.000	-0.000	0.2556	1.000	
9	-0.000	-0.000	0.2557	1.000	
10	-0.000	-0.000	0.2558	1.000	
11	-0.000	-0.000	0.2561	1.000	
12	-0.000	-0.000	0.2567	1.000	
13	-0.000	-0.000	0.2571	1.000	
14	-0.000	-0.000	0.2573	1.000	
15	-0.000	-0.000	0.2579	1.000	
16	-0.000	-0.000	0.2585	1.000	
17	-0.000	-0.000	0.2590	1.000	
18	-0.001	-0.000	0.2597	1.000	
19	-0.000	-0.000	0.2603	1.000	
20	-0.000	-0.000	0.2608	1.000	
21	-0.000	-0.000	0.2614	1.000	
22	-0.000	-0.000	0.2620	1.000	
23	-0.000	-0.000	0.2621	1.000	
24	-0.000	-0.000	0.2627	1.000	
25	-0.000	-0.000	0.2633	1.000	
26	-0.000	-0.000	0.2637	1.000	
27	-0.000	-0.000	0.2644	1.000	
28	-0.000	-0.000	0.2650	1.000	
29	-0.000	-0.000	0.2654	1.000	
30	-0.000	-0.000	0.2661	1.000	
31	-0.001	-0.000	0.2667	1.000	
32	-0.000	-0.000	0.2671	1.000	
33	-0.000	-0.000	0.2675	1.000	
34	-0.001	-0.000	0.2682	1.000	
35	-0.000	-0.000	0.2688	1.000	
36	-0.000	-0.000	0.2694	1.000	

FAC e FACP dos resíduos
quadráticos padronizados

Heteroskedasticity Test: ARCH

F-statistic	0.007166	Prob. F(36,2412)	1.0000
Obs*R-squared	0.261915	Prob. Chi-Square(36)	1.0000

(2) Modelo ARIMA-GARCH para a componente de detalhe $\left(\tilde{f}_{D_2}(t)\right)_{t=1}^{2506}$, onde o modelo ARIMA é dado por:

$$\begin{aligned} \hat{f}_{D_2}(t) = & \left[-\tilde{f}_{D_2}(t-2) \times 0.910567555981 - \tilde{f}_{D_2}(t-4) \times 0.454540567397 - \right. \\ & \tilde{f}_{D_2}(t-6) \times 0.434763130049 - \tilde{f}_{D_2}(t-8) \times 0.0625496805495 - \tilde{f}_{D_2}(t- \\ & 12) \times 0.0061730733711 + \tilde{f}_{D_2}(t-16) \times 0.0113950825853 + \tilde{f}_{D_2}(t- \\ & 20) \times 0.00350062813297 + \tilde{f}_{D_2}(t-28) \times 0.0255011965397 \left. \right] + \left[\hat{e}_{D_2}(t- \right. \\ & \left. 1) \times 0.975498319416 - \hat{e}_{D_2}(t-10) \times 0.0194826470364 \right] + \hat{\sigma}_{D_2,t}^2 \times 0. \end{aligned}$$

e o modelo GARCH (1,1), com distribuição do erro generalizada (GED), é dado por:

$$\begin{aligned} \hat{\sigma}_{D_2,t}^2 = & 0.00173751461567 - \left(\hat{e}_{A_2}(t-1) \right)^2 \times 0.000398540442848 - \\ & \hat{\sigma}_{D_2,t-1}^2 \times 0.968578671814. \end{aligned}$$

Dependent Variable: D2_F05X

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(2)	-0.951362	0.001307	-727.8200	0.0000
AR(4)	-0.461057	0.001507	-306.0015	0.0000
AR(6)	-0.452803	0.001436	-315.3974	0.0000
AR(8)	-0.027223	0.000582	-46.78741	0.0000
AR(12)	-0.002890	0.000108	-26.87308	0.0000
AR(16)	0.005812	0.000141	41.26588	0.0000
AR(20)	0.001524	6.40E-05	23.82591	0.0000
AR(28)	0.017024	0.000435	39.16344	0.0000
MA(1)	0.991330	0.000229	4319.569	0.0000
MA(10)	-0.006477	0.000215	-30.17886	0.0000
Variance Equation				
C	0.001977	6.02E-06	328.2736	0.0000
RESID(-1)^2	-0.000734	3.23E-05	-22.71988	0.0000
GARCH(-1)	-0.900873	0.070106	-12.85020	0.0000
GED PARAMETER	0.216378	0.004249	50.92055	0.0000
R-squared	0.680213	Mean dependent var	8.07E-06	
Adjusted R-squared	0.679047	S.D. dependent var	0.068509	
S.E. of regression	0.038812	Akaike info criterion	-7.322129	
Sum squared resid	3.716242	Schwarz criterion	-7.289264	
Log likelihood	9082.457	Hannan-Quinn criter.	-7.310192	
Durbin-Watson stat	1.996576			

		1	0.005	0.005	0.0599
		2	0.000	0.000	0.0602
		3	-0.001	-0.001	0.0634
		4	-0.024	-0.024	1.4371
		5	0.001	0.001	1.4380
		6	0.002	0.002	1.4525
		7	-0.001	-0.001	1.4543
		8	-0.023	-0.024	2.8108
		9	0.000	0.001	2.8113
		10	-0.003	-0.002	2.8281
		11	-0.001	-0.001	2.8290
		12	-0.040	-0.041	6.8074
		13	-0.001	-0.000	6.8081
		14	0.002	0.002	6.8162
		15	-0.000	-0.001	6.8166
		16	0.015	0.012	7.3582
		17	-0.000	-0.000	7.3582
		18	0.000	0.001	7.3588
		19	0.000	-0.000	7.3588
		20	0.014	0.012	7.8226
		21	-0.000	-0.000	7.8227
		22	0.000	0.000	7.8230
		23	0.000	0.000	7.8232
		24	-0.012	-0.012	8.1611
		25	0.000	0.000	8.1611
		26	-0.000	-0.000	8.1612
		27	-0.000	-0.000	8.1612
		28	0.010	0.011	8.4031
		29	0.000	0.000	8.4037
		30	0.001	0.002	8.4093
		31	0.000	0.000	8.4094
		32	-0.003	-0.002	8.4298
		33	0.000	0.000	8.4299
		34	0.001	0.001	8.4321
		35	0.000	0.000	8.4322
		36	0.003	0.003	8.4629

FAC e FACP dos resíduos
ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*	
		1	-0.000	-0.000	2E-06	0.999
		2	-0.000	-0.000	1E-05	1.000
		3	-0.000	-0.000	2E-05	1.000
		4	0.001	0.001	0.0007	1.000
		5	-0.000	-0.000	0.0007	1.000
		6	-0.000	-0.000	0.0007	1.000
		7	-0.000	-0.000	0.0007	1.000
		8	0.001	0.001	0.0014	1.000
		9	-0.000	-0.000	0.0014	1.000
		10	-0.000	-0.000	0.0014	1.000
		11	-0.000	-0.000	0.0014	1.000
		12	0.002	0.002	0.0114	1.000
		13	-0.000	-0.000	0.0114	1.000
		14	-0.000	-0.000	0.0114	1.000
		15	-0.000	-0.000	0.0114	1.000
		16	0.000	0.000	0.0115	1.000
		17	-0.000	-0.000	0.0115	1.000
		18	-0.000	-0.000	0.0116	1.000
		19	-0.000	-0.000	0.0116	1.000
		20	0.000	0.000	0.0116	1.000
		21	-0.000	-0.000	0.0116	1.000
		22	-0.000	-0.000	0.0116	1.000
		23	-0.000	-0.000	0.0116	1.000
		24	0.000	0.000	0.0116	1.000
		25	-0.000	-0.000	0.0117	1.000
		26	-0.000	-0.000	0.0117	1.000
		27	-0.000	-0.000	0.0117	1.000
		28	0.000	0.000	0.0117	1.000
		29	-0.000	-0.000	0.0117	1.000
		30	-0.000	-0.000	0.0117	1.000
		31	-0.000	-0.000	0.0117	1.000
		32	0.000	0.000	0.0117	1.000
		33	-0.000	-0.000	0.0118	1.000
		34	-0.000	-0.000	0.0118	1.000
		35	-0.000	-0.000	0.0118	1.000
		36	-0.000	-0.000	0.0118	1.000

FAC e FACP dos resíduos
quadráticos padronizados

Heteroskedasticity Test: ARCH

F-statistic	0.000436	Prob. F(1,2474)	0.9833
Obs*R-squared	0.000437	Prob. Chi-Square(1)	0.9833

(3) Modelo ARIMA-GARCH para a componente de detalhe $\left(\tilde{f}_{D_3}(t)\right)_{t=1}^{2506}$, onde o modelo ARIMA é dado por:

$$\begin{aligned} \hat{f}_{D_3}(t) = & \left[-\tilde{f}_{D_3}(t-1) \times 0.961642062671 - \tilde{f}_{D_3}(t-2) \times 0.0456818533945 - \right. \\ & \tilde{f}_{D_3}(t-3) \times 0.0808462767147 + \tilde{f}_{D_3}(t-4) \times 0.00321476100108 - \\ & \left. \tilde{f}_{D_3}(t-8) \times 1.27961205305 \times 10^{-5} - \tilde{f}_{D_3}(t-18) \right] \times 8.24028070289 \times 10^{-6} + \\ & \left[-\hat{e}_{D_3}(t-1) \times 0.03835802031 - \hat{e}_{D_3}(t-2) \times 0.743445560408 + \right. \\ & \left. \hat{e}_{D_3}(t-3) \times 0.0285449308416 \right] + \hat{\sigma}_{D_3,t}^2 \times 4.95628401348 \times 10^{-6} \end{aligned}$$

e o modelo EGARCH (1,2), com ordem assimétrica igual a 1 e distribuição do erro generalizada (GED), é dado por:

$$\log(\hat{\sigma}_{D_3,t}^2) = -28.79582 - \left| \left(\frac{\hat{e}_{D_3}(t-1)}{\sqrt{\hat{\sigma}_{D_3,t-1}^2}} \right) \right| \times 0.079313 + \left(\frac{\hat{e}_{D_3}(t-1)}{\sqrt{\hat{\sigma}_{D_3,t-1}^2}} \right) \times 0.069759 - \log(\hat{\sigma}_{D_3,t-1}^2) \times 1.000640.$$

Dependent Variable: D1_F05X

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	1.30E-05	3.30E-06	3.945238	0.0001
AR(2)	-0.062044	3.87E-06	-16049.56	0.0000
AR(4)	0.003803	1.46E-06	2606.870	0.0000
AR(1)	-0.961494	1.24E-06	-778016.0	0.0000
AR(8)	-0.000103	1.52E-06	-68.08180	0.0000
AR(3)	-0.096898	2.18E-06	-44415.52	0.0000
AR(18)	-4.40E-05	7.02E-07	-62.71757	0.0000
MA(1)	-0.038496	1.11E-06	-34597.31	0.0000
MA(2)	-0.724079	0.000117	-6210.133	0.0000
MA(3)	0.028084	2.17E-06	12950.18	0.0000
Variance Equation				
C(11)	-28.79582	0.121286	-237.4211	0.0000
C(12)	-0.079313	0.006139	-12.91867	0.0000
C(13)	0.069759	0.005813	12.00130	0.0000
C(14)	-1.000640	0.000114	-8780.863	0.0000
GED PARAMETER	0.464876	0.006585	70.60000	0.0000
R-squared	0.484540	Mean dependent var		0.000000
Adjusted R-squared	0.482668	S.D. dependent var		0.032332
S.E. of regression	0.023255	Akaike info criterion		-13.15950
Sum squared resid	1.339551	Schwarz criterion		-13.12441
Log likelihood	16378.84	Hannan-Quinn criter.		-13.14676
Durbin-Watson stat	1.999558			

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	-0.015	-0.015	0.5564		
2	0.027	0.027	2.3594		
3	-0.003	-0.002	2.3787		
4	0.002	0.001	2.3872		
5	0.005	0.005	2.4384		
6	-0.002	-0.002	2.4533		
7	0.007	0.007	2.5848		
8	-0.003	-0.002	2.6031		
9	0.001	0.001	2.6058	0.106	
10	0.015	0.015	3.1473	0.207	
11	0.011	0.012	3.4636	0.326	
12	-0.003	-0.004	3.4943	0.479	
13	-0.004	-0.004	3.5281	0.619	
14	-0.001	-0.001	3.5295	0.740	
15	-0.004	-0.004	3.5632	0.828	
16	0.028	0.028	5.5191	0.701	
17	-0.005	-0.005	5.5927	0.780	
18	0.041	0.039	9.8359	0.455	
19	0.009	0.011	10.062	0.525	
20	0.042	0.040	14.416	0.275	
21	0.002	0.003	14.428	0.344	
22	0.016	0.014	15.039	0.375	
23	-0.002	-0.002	15.047	0.448	
24	0.014	0.014	15.561	0.484	
25	-0.007	-0.007	15.682	0.546	
26	0.034	0.033	18.679	0.412	
27	0.002	0.003	18.692	0.477	
28	0.033	0.031	21.410	0.373	
29	-0.003	-0.003	21.431	0.433	
30	0.055	0.053	29.020	0.144	
31	0.000	0.001	29.020	0.180	
32	0.038	0.036	32.658	0.112	
33	0.009	0.010	32.847	0.135	
34	-0.002	-0.004	32.854	0.166	
35	-0.009	-0.010	33.050	0.195	
36	-0.004	-0.007	33.087	0.233	

FAC e FACP dos resíduos
ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	0.006	0.006	0.0817	0.775	
2	0.001	0.001	0.0835	0.959	
3	-0.001	-0.001	0.0878	0.993	
4	-0.000	-0.000	0.0880	0.999	
5	-0.003	-0.003	0.1039	1.000	
6	0.001	0.001	0.1083	1.000	
7	-0.003	-0.003	0.1255	1.000	
8	-0.000	-0.000	0.1260	1.000	
9	-0.003	-0.003	0.1446	1.000	
10	0.000	0.001	0.1452	1.000	
11	-0.003	-0.003	0.1620	1.000	
12	0.000	0.000	0.1621	1.000	
13	-0.003	-0.003	0.1820	1.000	
14	0.000	0.000	0.1821	1.000	
15	-0.003	-0.003	0.2000	1.000	
16	0.002	0.002	0.2086	1.000	
17	-0.003	-0.003	0.2287	1.000	
18	-0.000	-0.000	0.2293	1.000	
19	-0.003	-0.003	0.2481	1.000	
20	0.000	0.000	0.2482	1.000	
21	-0.003	-0.003	0.2692	1.000	
22	-0.000	-0.000	0.2694	1.000	
23	-0.003	-0.003	0.2908	1.000	
24	-0.001	-0.001	0.2924	1.000	
25	-0.003	-0.003	0.3133	1.000	
26	-0.000	-0.000	0.3133	1.000	
27	-0.003	-0.003	0.3338	1.000	
28	0.000	0.000	0.3340	1.000	
29	-0.003	-0.003	0.3530	1.000	
30	0.003	0.003	0.3706	1.000	
31	-0.003	-0.003	0.3910	1.000	
32	-0.001	-0.001	0.3920	1.000	
33	-0.003	-0.003	0.4124	1.000	
34	0.001	0.001	0.4165	1.000	
35	-0.003	-0.003	0.4363	1.000	
36	0.002	0.002	0.4468	1.000	

FAC e FACP dos resíduos
quadráticos padronizados

Apêndice I D: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX-GARCH Neural (F05x).

Na Figura 59, podemos comparar a evolução temporal das previsões fora da amostra para WARIMAX-GARCH Neural e ARIMA-GARCH em relação às observações de fora da amostra

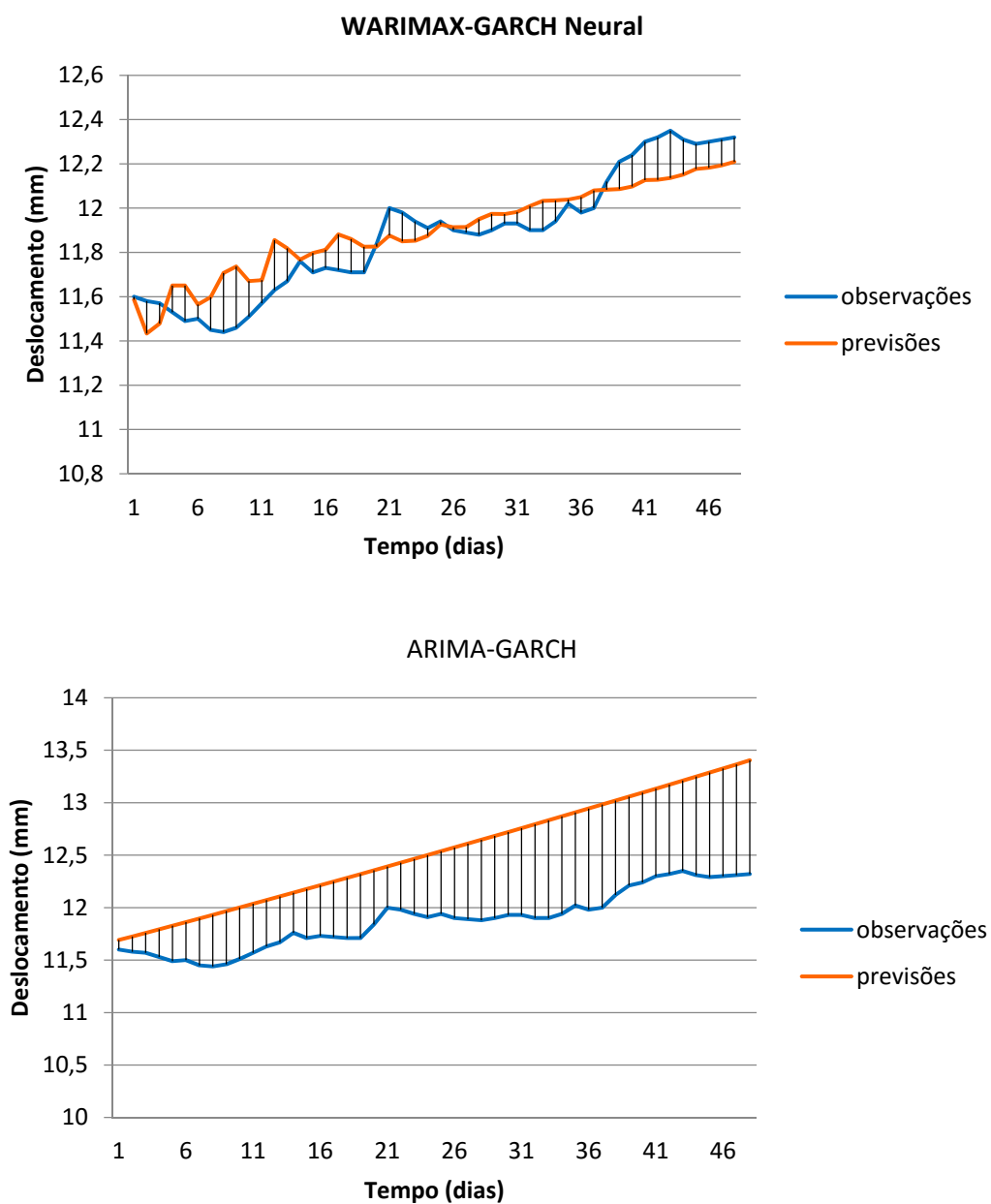


Figura 59 – Evolução temporal do erro APE para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA

Dependent Variable: F05X

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	2.083062	0.153600	13.56160	0.0000
D2_F05XC(-1)	1.435834	0.004214	340.7318	0.0000
D2_F05XC(-2)	-0.640943	0.002357	-271.9087	0.0000
A2_F05XC(-1)	1.592509	5.40E-05	29476.81	0.0000

A2_F05XC(-2)	-0.593229	8.39E-05	-7074.538	0.0000
AR(1)	-0.500193	0.002726	-183.4763	0.0000
AR(3)	0.174918	0.010881	16.07511	0.0000
AR(4)	0.850507	0.004452	191.0283	0.0000
AR(5)	0.274827	0.009907	27.74105	0.0000
MA(2)	-0.035774	0.004287	-8.344586	0.0000
MA(3)	-0.151000	0.010679	-14.13968	0.0000
MA(4)	-0.729599	0.008951	-81.50948	0.0000
MA(5)	0.123217	0.006980	17.65364	0.0000
Variance Equation				
C	0.000379	7.00E-05	5.414488	0.0000
RESID(-1)^2	0.007705	0.002083	3.698137	0.0002
GARCH(-1)	0.892784	0.017407	51.28807	0.0000
GED PARAMETER	0.503908	0.013619	36.99975	0.0000
R-squared	0.997374	Mean dependent var	11.37928	
Adjusted R-squared	0.997361	S.D. dependent var	1.362144	
S.E. of regression	0.069970	Akaike info criterion	-3.591017	
Sum squared resid	12.16595	Schwarz criterion	-3.551387	
Log likelihood	4502.180	Hannan-Quinn criter.	-3.576629	
Durbin-Watson stat	2.015345			

Apêndice II A: Teste Dikey-Fuller Aumentado (série temporal F05y).

Null Hypothesis: F05Y has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.650913	0.0049
Test critical values: 1% level	-3.433368	
5% level	-2.862759	
10% level	-2.567465	

Null Hypothesis: D(F05Y) has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.549567	0.0000
Test critical values: 1% level	-3.433368	
5% level	-2.862759	
10% level	-2.567465	

Apêndice II B: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (série temporal F05y).

O modelo ARIMA (3,1,1) é dado por:

$$\nabla(\hat{f}(t)) = [\tilde{f}(t-1) \times 1.066177 - \tilde{f}(t-3) \times 0.079496] - [\hat{e}(t-1) \times 0.948427]$$

onde ∇ denota o operador diferença de primeira ordem não sazonal. Por sua vez, o modelo GARCH (2,1) é dado por:

$$\hat{\sigma}_t^2 = 4.14 \times 10^{-5} + (\hat{e}(t-1))^2 \times 0.147979 + \hat{\sigma}_{A_2t-1}^2 \times 0.260716 + \hat{\sigma}_{A_2t-2}^2 \times 0.550662.$$

Dependent Variable: D(F05Y)

Method: ML - ARCH (Marquardt) - Normal distribution

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	1.066177	0.017368	61.38819	0.0000
AR(3)	-0.079496	0.016283	-4.881999	0.0000
MA(1)	-0.948427	0.007577	-125.1715	0.0000

Variance Equation

C	4.14E-05	8.47E-06	4.881298	0.0000
RESID(-1)^2	0.147979	0.014560	10.16318	0.0000
GARCH(-1)	0.260716	0.094555	2.757297	0.0058
GARCH(-2)	0.550662	0.091880	5.993262	0.0000

R-squared	0.265749	Mean dependent var	-0.001665
Adjusted R-squared	0.264652	S.D. dependent var	0.035628
S.E. of regression	0.030552	Akaike info criterion	-4.277041
Sum squared resid	1.874340	Schwarz criterion	-4.254747
Log likelihood	4310.703	Hannan-Quinn criter.	-4.268858
Durbin-Watson stat	1.967537		

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 0.015	0.015	0.4890	
		2 0.022	0.021	1.4021	
		3 -0.018	-0.019	2.0710	
		4 -0.042	-0.042	5.6430	
		5 0.023	0.025	6.7327	0.019
		6 -0.005	-0.005	6.7898	0.034
		7 0.015	0.013	7.2668	0.064
		8 -0.009	-0.011	7.4443	0.114
		9 -0.002	-0.001	7.4547	0.189
		10 0.017	0.017	8.0080	0.238
		11 0.028	0.028	9.5659	0.215
		12 0.028	0.025	11.118	0.195
		13 0.021	0.021	12.047	0.211
		14 0.007	0.008	12.154	0.275
		15 -0.006	-0.004	12.223	0.347
		16 0.006	0.008	12.308	0.421
		17 0.026	0.026	13.633	0.400
		18 0.001	-0.001	13.635	0.477
		19 0.036	0.035	16.282	0.364
		20 -0.036	-0.035	18.881	0.275
		21 0.041	0.042	22.321	0.173
		22 -0.005	-0.006	22.379	0.216
		23 -0.002	-0.004	22.389	0.265
		24 -0.016	-0.022	22.944	0.292
		25 -0.018	-0.012	23.570	0.314
		26 0.025	0.022	24.813	0.306
		27 -0.010	-0.010	25.025	0.349
		28 0.036	0.030	27.617	0.277
		29 0.013	0.011	27.977	0.309
		30 0.010	0.009	28.188	0.349
		31 0.023	0.020	29.224	0.350
		32 0.018	0.020	29.909	0.368
		33 0.010	0.007	30.134	0.407
		34 0.018	0.018	30.779	0.426
		35 0.010	0.011	30.988	0.467
		36 -0.001	-0.000	30.989	0.518

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 0.061	0.061	7.5727	0.013
		2 -0.022	-0.026	8.5789	0.014
		3 -0.005	-0.002	8.6300	0.035
		4 -0.001	-0.001	8.6330	0.071
		5 -0.042	-0.042	12.245	0.032
		6 -0.000	0.005	12.246	0.057
		7 -0.018	-0.020	12.871	0.075
		8 -0.028	-0.026	14.411	0.072
		9 -0.038	-0.036	17.327	0.044
		10 -0.024	-0.023	18.463	0.048
		11 -0.051	-0.051	23.814	0.014
		12 -0.016	-0.014	24.358	0.018
		13 0.011	0.008	24.606	0.026
		14 0.009	0.003	24.763	0.037
		15 0.015	0.012	25.243	0.047
		16 0.027	0.020	26.766	0.044
		17 0.015	0.009	27.217	0.055
		18 -0.002	-0.005	27.224	0.075
		19 0.013	0.010	27.548	0.093
		20 -0.021	-0.026	28.460	0.099
		21 0.010	0.013	28.681	0.122
		22 0.019	0.016	29.411	0.134
		23 -0.001	-0.002	29.412	0.167
		24 -0.017	-0.011	29.986	0.185
		25 0.004	0.008	30.027	0.223
		26 0.027	0.030	31.505	0.210
		27 0.017	0.018	32.063	0.230
		28 0.005	0.006	32.109	0.270
		29 0.005	0.004	32.154	0.313
		30 0.061	0.064	39.877	0.107
		31 -0.017	-0.023	40.461	0.119
		32 -0.018	-0.012	41.140	0.129
		33 0.013	0.018	41.493	0.147
		34 0.028	0.028	43.094	0.136
		35 0.001	0.005	43.095	0.164
		36 0.017	0.020	43.684	0.177

FAC e FACP dos resíduos
ordinários padronizados

FAC e FACP dos resíduos
quadráticos padronizados

Apêndice II C: Configurações da RNA (série temporal F05y).

Tabela 8 – Configurações das RNA-MLPs para as componentes wavelet.

<i>Parâmetros da arquitetura neural</i>	<i>Componente de aproximação nível 2</i>
Transformação	premnmx
Tamanho de janela de entrada	2
Número de camadas escondidas	1
Número de neurônio na camada escondida	3
Função de ativação da camada oculta	Tangente hiperbólica
Função de ativação da camada de saída	Linear
Algoritmo de treinamento	Levenberg-Marquardt (Millar, 2011)

Apêndice II D: Modelo WARIMAX-GARCH e seus resultados estatísticos para validação (F05y).

(1) Modelo ARIMA-GARCH para a componente de aproximação $\left(\tilde{f}_{A_2}(t)\right)_{t=1}^{2016}$,

onde o modelo ARIMA é dado por:

$$\nabla^3 \left(\hat{f}_{A_2}(t) \right) = \left[\tilde{f}_{A_2}(t-1) \times 2.352910 - \tilde{f}_{A_2}(t-2) \times 1.904162 + \tilde{f}_{A_2}(t-4) \times 0.489154 - \tilde{f}_{A_2}(t-9) \times 0.882663 + \tilde{f}_{A_2}(t-10) \times 1.054069 + \tilde{f}_{A_2}(t-12) \times 0.443656 + \tilde{f}_{A_2}(t-15) \times 0.140106 - \tilde{f}_{A_2}(t-18) \times 0.085887 \right] + \tilde{f}_{A_2}(t-20) \times 0.043912 + \left[\hat{e}_{A_2}(t-1) \times 0.281448 + \hat{e}_{A_2}(t-2) \times 0.770833 - \hat{e}_{A_2}(t-4) \times 0.590304 - \hat{e}_{A_2}(t-5) \times 0.681541 - \hat{e}_{A_2}(t-6) \times 0.632482 + \hat{e}_{A_2}(t-8) \times 0.073315 + \hat{e}_{A_2}(t-9) \times 0.627362 + \hat{e}_{A_2}(t-10) \times 0.113089 - \hat{e}_{A_2}(t-12) \times 0.132023 - \hat{e}_{A_2}(t-13) \times 0.340584 - \hat{e}_{A_2}(t-15) \times 0.132267 + 0 \times \hat{\sigma}_{A_2,t}^2 \right]$$

e o modelo GARCH (1,1), com distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{A_2,t}^2 = 1.22 \times 10^{-9} + 0.045825 \times \left(\hat{e}_{A_2}(t-1) \right)^2 + 0.806679 \times \hat{\sigma}_{A_2,t-1}^2.$$

Dependent Variable: D(A2_F05Y,3)

GARCH = C(21) + C(22)*RESID(-1)^2 + C(23)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	2.352910	0.014155	166.2298	0.0000
AR(2)	-1.904162	0.017040	-111.7497	0.0000
AR(4)	0.489154	0.012203	40.08560	0.0000
AR(9)	-0.882663	0.018898	-46.70567	0.0000
AR(20)	0.043912	0.004651	9.441716	0.0000
AR(18)	-0.085887	0.004503	-19.07278	0.0000
AR(15)	0.140106	0.008521	16.44181	0.0000
AR(12)	-0.443656	0.012768	-34.74787	0.0000
AR(10)	1.054069	0.026691	39.49225	0.0000
MA(1)	0.281448	0.020028	14.05237	0.0000
MA(2)	0.770833	0.023940	32.19830	0.0000
MA(4)	-0.590304	0.036055	-16.37227	0.0000
MA(5)	-0.681541	0.027457	-24.82206	0.0000
MA(6)	-0.632482	0.030169	-20.96476	0.0000
MA(8)	0.073315	0.031589	2.320904	0.0203
MA(9)	0.627362	0.025888	24.23381	0.0000
MA(10)	0.113089	0.034449	3.282813	0.0010
MA(12)	-0.132023	0.021316	-6.193580	0.0000
MA(13)	-0.340584	0.027050	-12.59109	0.0000
MA(15)	-0.132267	0.021937	-6.029489	0.0000
Variance Equation				
C	1.22E-09	7.35E-10	1.656864	0.0975
RESID(-1)^2	0.045825	0.020145	2.274737	0.0229
GARCH(-1)	0.806679	0.102560	7.865453	0.0000
GED PARAMETER	1.908215	0.093837	20.33535	0.0000
R-squared	0.999594	Mean dependent var	3.51E-06	
Adjusted R-squared	0.999590	S.D. dependent var	0.004517	
S.E. of regression	9.14E-05	Akaike info criterion	-15.75975	
Sum squared resid	1.65E-05	Schwarz criterion	-15.69234	
Log likelihood	15728.59	Hannan-Quinn criter.	-15.73499	
Durbin-Watson stat	1.797000			

Heteroskedasticity Test: ARCH

F-statistic	0.007166	Prob. F(36,2412)	1.0000
Obs*R-squared	0.261915	Prob. Chi-Square(36)	1.0000

(2) Modelo ARIMA-GARCH para a componente de detalhe $\left(\tilde{f}_{D_2}(t)\right)_{t=1}^{2016}$ é dado

por:

$$\begin{aligned} \hat{f}_{D_2}(t) = & \left[-\tilde{f}_{D_2}(t-2) \times 0.910567555981 - \tilde{f}_{D_2}(t-4) \times 0.454540567397 - \right. \\ & \tilde{f}_{D_2}(t-6) \times 0.434763130049 - \tilde{f}_{D_2}(t-8) \times 0.0625496805495 - \tilde{f}_{D_2}(t- \\ & 12) \times 0.0061730733711 + \tilde{f}_{D_2}(t-16) \times 0.0113950825853 + \tilde{f}_{D_2}(t- \\ & 20) \times 0.00350062813297 + \tilde{f}_{D_2}(t-28) \times 0.0255011965397 \left. \right] + \left[\hat{e}_{D_2}(t- \right. \\ & \left. 1) \times 0.975498319416 - \hat{e}_{D_2}(t-10) \times 0.0194826470364 \right] + \hat{\sigma}_{D_2,t}^2 \times 0. \end{aligned}$$

Dependent Variable: D2_F05Y

Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-2.164983	0.048642	-44.50844	0.0000
AR(2)	-3.438877	0.116853	-29.42905	0.0000
AR(3)	-4.056498	0.195435	-20.75629	0.0000
AR(4)	-3.917835	0.238898	-16.39959	0.0000
AR(5)	-2.730596	0.230378	-11.85268	0.0000
AR(6)	-1.695662	0.165673	-10.23502	0.0000
AR(7)	-0.641666	0.092072	-6.969159	0.0000
AR(8)	-0.178234	0.032805	-5.433133	0.0000
MA(1)	-2.858902	0.045290	-63.12440	0.0000
MA(2)	3.398204	0.118884	28.58415	0.0000
MA(3)	-1.185617	0.135315	-8.761888	0.0000
MA(4)	-0.957121	0.066306	-14.43482	0.0000
MA(6)	2.529027	0.081171	31.15683	0.0000
MA(7)	-3.180853	0.139358	-22.82506	0.0000
MA(8)	1.767870	0.103032	17.15846	0.0000
MA(9)	-0.409377	0.032978	-12.41372	0.0000
R-squared	0.997370	Mean dependent var	1.80E-06	
Adjusted R-squared	0.997350	S.D. dependent var	0.011162	
S.E. of regression	0.000575	Akaike info criterion	-12.07801	
Sum squared resid	0.000658	Schwarz criterion	-12.03335	
Log likelihood	12142.32	Hannan-Quinn criter.	-12.06162	
Durbin-Watson stat	1.998269			

Heteroskedasticity Test: ARCH

F-statistic	0.017878	Prob. F(1,1990)	0.8936
Obs*R-squared	0.017896	Prob. Chi-Square(1)	0.8936

(3) Modelo ARIMA-GARCH para a componente de detalhe $\left(\tilde{f}_{D_3}(t)\right)_{t=1}^{2016}$, onde o

modelo ARIMA é dado por:

$$\begin{aligned} \hat{f}_{D_3}(t) = & [\tilde{f}_{D_3}(t-1) \times 0.880721 - \tilde{f}_{D_3}(t-2) \times 2.147480 - \tilde{f}_{D_3}(t-3) \times 0.269246 + \tilde{f}_{D_3}(t-4) \times 0.895915 - \tilde{f}_{D_3}(t-5) \times 1.410807 - \tilde{f}_{D_3}(t-7) \times 0.928008 + \tilde{f}_{D_3}(t-8) \times 0.527949 + \tilde{f}_{D_3}(t-10) \times 0.353120] + \\ & [\hat{e}_{D_3}(t-1) \times 0.735844 - \hat{e}_{D_3}(t-2) \times 1.718650 - \hat{e}_{D_3}(t-3) \times 1.961380 + \hat{e}_{D_3}(t-4) \times 0.524908 + \hat{e}_{D_3}(t-5) \times 1.065598 + \hat{e}_{D_3}(t-7) \times 0.451835 + \hat{e}_{D_3}(t-8) \times 0.774244 - \hat{e}_{D_3}(t-9) \times 0.327392 - \hat{e}_{D_3}(t-10) \times 0.544504] \end{aligned}$$

e o modelo GARCH (1,1), com ordem assimétrica igual a 1 e distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{D_2,t}^2 = 4.19^{-9} - \left(\hat{e}_{A_2}(t-1)\right)^2 \times 0.182185 - \hat{\sigma}_{D_2,t-1}^2 \times 0.796036.$$

Dependent Variable: D(D3_F05Y)

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

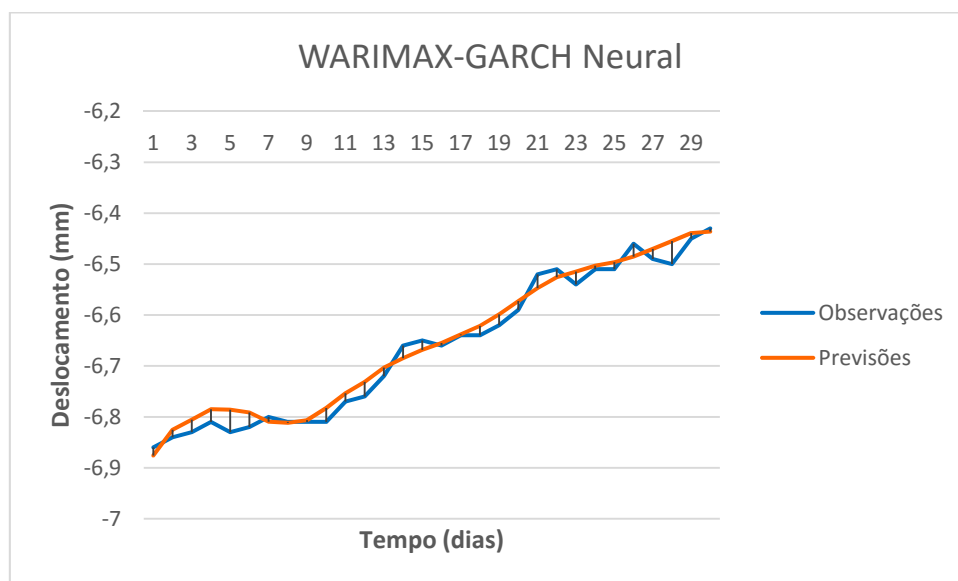
Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.880721	0.024456	36.01220	0.0000
AR(2)	-2.147480	0.032946	-65.18192	0.0000
AR(3)	0.269246	0.065579	4.105678	0.0000
AR(4)	-0.895915	0.042313	-21.17360	0.0000
AR(5)	-1.410807	0.051693	-27.29182	0.0000
AR(7)	-0.928008	0.021393	-43.37978	0.0000
AR(8)	-0.527949	0.018298	-28.85296	0.0000
AR(10)	-0.353120	0.012027	-29.36060	0.0000
MA(1)	0.735844	0.004003	183.8265	0.0000
MA(2)	-1.718650	0.001306	-1315.532	0.0000
MA(3)	-1.961380	0.001011	-1940.150	0.0000
MA(4)	0.524908	0.000175	3001.364	0.0000
MA(5)	1.065598	0.000236	4522.475	0.0000
MA(7)	0.451835	0.001578	286.2977	0.0000
MA(8)	0.774244	0.002021	383.1531	0.0000
MA(9)	-0.327392	0.003251	-100.6917	0.0000
MA(10)	-0.544504	0.003646	-149.3547	0.0000

Variance Equation

C	4.19E-09	1.08E-09	3.896222	0.0001
RESID(-1)^2	0.182185	0.025711	7.085965	0.0000
GARCH(-1)	0.796036	0.022861	34.82139	0.0000
GED PARAMETER	2.020774	0.105281	19.19412	0.0000
R-squared	0.999444	Mean dependent var	3.98E-06	
Adjusted R-squared	0.999440	S.D. dependent var	0.015374	
S.E. of regression	0.000364	Akaike info criterion	-13.21117	
Sum squared resid	0.000263	Schwarz criterion	-13.15248	
Log likelihood	13265.20	Hannan-Quinn criter.	-13.18962	
Durbin-Watson stat	1.622453			

Apêndice II E: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX-GARCH Neural (F05y).

Na Figura 60, pode-se comparar a evolução temporal das previsões fora da amostra para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA em relação às observações de fora da amostra



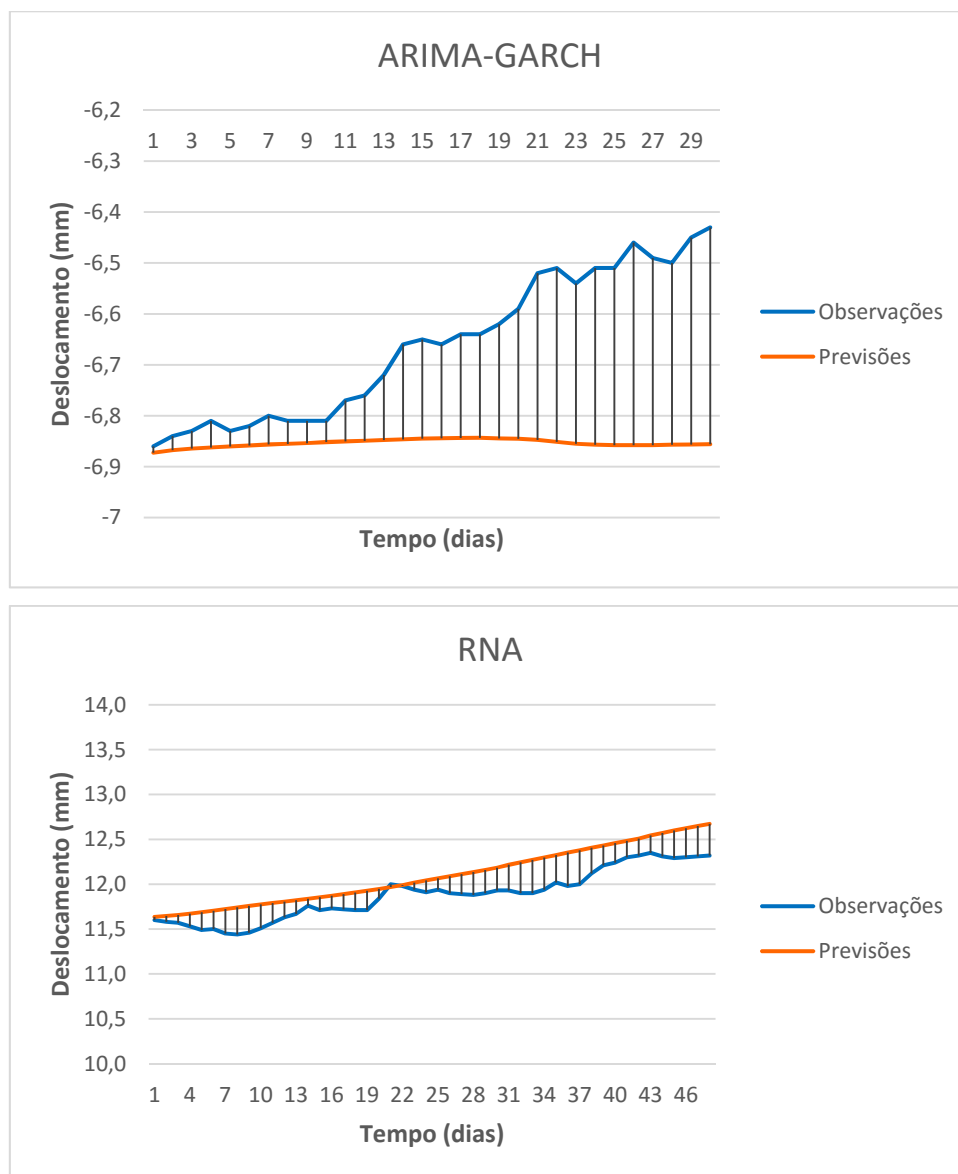


Figura 60 – Evolução temporal do erro APE para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA

Dependent Variable: F05Y

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
A2_F05YF(-1)	1.984882	0.004073	487.3464	0.0000
A2_F05YF(-2)	-0.984894	0.004074	-241.7612	0.0000
D2_F05Y(-1)	0.251687	0.010671	-90.37913	0.0000
D3_F05Y(-1)	0.454555	0.020270	39.65699	0.0000
AR(1)	0.844260	0.027126	31.12363	0.0000
AR(2)	-0.377094	0.029262	-12.88696	0.0000
AR(3)	-0.107103	0.024826	-4.314131	0.0000
MA(1)	-0.868101	0.017349	-50.03686	0.0000

Variance Equation				
C	1.34E-05	3.72E-06	3.590341	0.0003
RESID(-1)^2	0.108030	0.017042	6.339191	0.0000
GARCH(-1)	0.843740	0.024216	34.84199	0.0000
GED PARAMETER	1.448351	0.055848	25.93392	0.0000
R-squared	0.999784	Mean dependent var	-5.104764	
Adjusted R-squared	0.999783	S.D. dependent var	1.122720	
S.E. of regression	0.016536	Akaike info criterion	-5.521290	
Sum squared resid	0.548252	Schwarz criterion	-5.493412	
Log likelihood	5561.657	Hannan-Quinn criter.	-5.511057	
Durbin-Watson stat	2.019100			

Heteroskedasticity Test: ARCH

F-statistic	1.735737	Prob. F(1,2008)	0.1878
Obs*R-squared	1.735965	Prob. Chi-Square(1)	0.1877

Apêndice III A: Teste Dikey-Fuller Aumentado (F13x).

Null Hypothesis: F13X has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.513505	0.5268
Test critical values: 1% level	-3.433387	
5% level	-2.862768	
10% level	-2.567470	

Null Hypothesis: D(F13X) has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-34.58909	0.0000
Test critical values: 1% level	-3.433387	
5% level	-2.862768	
10% level	-2.567470	

*MacKinnon (1996) one-sided p-values.

Apêndice III B: Configurações da RNA (série temporal F13x).

Tabela 9 – Configurações das RNA-MLPs para as componentes *wavelet*.

Parâmetros da arquitetura neural	Componente de aproximação nível 2
Transformação	premnmx
Tamanho de janela de entrada	3
Número de camadas escondidas	1
Número de neurônio na camada escondida	5
Função de ativação da camada oculta	Tangente hiperbólica
Função de ativação da camada de saída	Linear
Algoritmo de treinamento	Levenberg-Marquardt (Millar, 2011)

Apêndice III C: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (F13x).

O modelo ARIMA (18,1,1), com a transformação logaritmo, é dado por:

$$\nabla(\hat{f}(t)) = [\tilde{f}(t-1) \times 1.352494 - \tilde{f}(t-2) \times 0.369022] + [-\hat{e}(t-1) \times 0.964281 - \hat{e}(t-1) \times 0.031454]$$

onde ∇ denota o operador diferença de primeira ordem não sazonal. Por sua vez, o modelo GARCH (1,1) é dado por:

$$\hat{\sigma}_t^2 = 2.05 \times 10^{-5} + (\hat{e}(t-1))^2 \times 0.180137 + \hat{\sigma}_{A_2t-1}^2 \times 0.724233$$

Dependent Variable: D(LOG(F13X))
Method: ML - ARCH

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	1.352494	0.025093	53.90005	0.0000
AR(2)	-0.369022	0.023561	-15.66239	0.0000
MA(1)	-0.964281	0.010669	-90.37937	0.0000
MA(18)	0.031454	0.007055	4.458748	0.0000
Variance Equation				
C	2.05E-06	2.43E-07	8.454114	0.0000
RESID(-1)^2	0.180137	0.014437	12.47703	0.0000
GARCH(-1)	0.724233	0.021422	33.80746	0.0000
R-squared	0.074383	Mean dependent var		-1.51E-05
Adjusted R-squared	0.072986	S.D. dependent var		0.005181
S.E. of regression	0.004989	Akaike info criterion		-8.125946
Sum squared resid	0.049452	Schwarz criterion		-8.106270
Log likelihood	8096.379	Hannan-Quinn criter.		-8.118720
Durbin-Watson stat	2.326397			

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	0.027	0.027	1.4512	0.228	
2	-0.009	-0.010	1.6154	0.446	
3	-0.023	-0.023	2.7118	0.438	
4	-0.020	-0.018	3.4776	0.481	
5	-0.010	-0.010	3.6825	0.596	
6	-0.001	-0.001	3.6834	0.719	
7	-0.027	-0.028	5.1793	0.638	
8	0.008	0.009	5.3173	0.723	
9	-0.015	-0.016	5.7579	0.764	
10	-0.013	-0.014	6.0999	0.807	
11	0.004	0.004	6.1314	0.864	
12	0.047	0.045	10.493	0.573	
13	0.091	0.088	26.965	0.013	
14	-0.012	-0.017	27.269	0.018	
15	0.012	0.017	27.550	0.025	
16	0.046	0.051	31.867	0.010	
17	0.017	0.019	32.478	0.013	
18	0.005	0.007	32.522	0.019	
19	0.024	0.029	33.687	0.020	
20	0.010	0.016	33.875	0.027	
21	0.045	0.046	37.903	0.013	
22	-0.012	-0.007	38.178	0.017	
23	-0.019	-0.012	38.914	0.020	
24	-0.011	-0.010	39.158	0.026	
25	-0.010	-0.015	39.350	0.034	
26	-0.002	-0.006	39.361	0.045	
27	0.029	0.029	41.039	0.041	
28	0.005	-0.002	41.095	0.053	
29	0.008	-0.003	41.220	0.066	
30	0.008	0.007	41.360	0.081	
31	-0.010	-0.013	41.567	0.097	
32	0.010	0.001	41.771	0.116	
33	0.020	0.012	42.566	0.123	
34	0.017	0.009	43.174	0.135	
35	0.008	0.009	43.306	0.158	
36	-0.001	0.001	43.307	0.188	

FAC e FACP dos resíduos
ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	0.027	0.027	1.4512	0.228	
2	-0.009	-0.010	1.6154	0.446	
3	-0.023	-0.023	2.7118	0.438	
4	-0.020	-0.018	3.4776	0.481	
5	-0.010	-0.010	3.6825	0.596	
6	-0.001	-0.001	3.6834	0.719	
7	-0.027	-0.028	5.1793	0.638	
8	0.008	0.009	5.3173	0.723	
9	-0.015	-0.016	5.7579	0.764	
10	-0.013	-0.014	6.0999	0.807	
11	0.004	0.004	6.1314	0.864	
12	0.047	0.045	10.493	0.573	
13	0.091	0.088	26.965	0.013	
14	-0.012	-0.017	27.269	0.018	
15	0.012	0.017	27.550	0.025	
16	0.046	0.051	31.867	0.010	
17	0.017	0.019	32.478	0.013	
18	0.005	0.007	32.522	0.019	
19	0.024	0.029	33.687	0.020	
20	0.010	0.016	33.875	0.027	
21	0.045	0.046	37.903	0.013	
22	-0.012	-0.007	38.178	0.017	
23	-0.019	-0.012	38.914	0.020	
24	-0.011	-0.010	39.158	0.026	
25	-0.010	-0.015	39.350	0.034	
26	-0.002	-0.006	39.361	0.045	
27	0.029	0.029	41.039	0.041	
28	0.005	-0.002	41.095	0.053	
29	0.008	-0.003	41.220	0.066	
30	0.008	0.007	41.360	0.081	
31	-0.010	-0.013	41.567	0.097	
32	0.010	0.001	41.771	0.116	
33	0.020	0.012	42.566	0.123	
34	0.017	0.009	43.174	0.135	
35	0.008	0.009	43.306	0.158	
36	-0.001	0.001	43.307	0.188	

FAC e FACP dos resíduos
quadráticos padronizados

Heteroskedasticity Test: ARCH

F-statistic	1.423157	Prob. F(1,1987)	0.2330
Obs*R-squared	1.423569	Prob. Chi-Square(1)	0.2328

Apêndice III D: Modelo WARIMAX-GARCH Neural e seus resultados estatísticos para validação (F13x).

(1) Modelo ARIMA-GARCH para a componente de aproximação $\left(\tilde{f}_{A_2}(t)\right)_{t=1}^{1994}$,

onde o modelo ARIMA é dado por:

$$\nabla^3 \left(\hat{f}_{A_2}(t) \right) = \left[\tilde{f}_{A_2}(t-1) \times 1.491942 - \tilde{f}_{A_2}(t-2) \times 1.214282 + \tilde{f}_{A_2}(t-3) \times 0.804555 + \tilde{f}_{A_2}(t-5) \times 1.184911 + \tilde{f}_{A_2}(t-6) \times 0.549937 \right] + \left[\hat{e}_{A_2}(t-1) \times 0.115114 - \hat{e}_{A_2}(t-2) \times 0.046410 - \hat{e}_{A_2}(t-3) \times 0.590304 - \hat{e}_{A_2}(t-4) \times 0.887668 + 0 \times \hat{\sigma}_{A_2,t}^2 \right]$$

e o modelo GARCH (1,1), com distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{A_2,t}^2 = 9.85 \times 10^{-9} + 0.265391 \times (\hat{e}_{A_2}(t-1))^2 + 0.703556 \times \hat{\sigma}_{A_2,t-1}^2.$$

Dependent Variable: D(LOG(A2_F13X),3)

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	1.491942	0.022134	67.40483	0.0000
AR(2)	-1.214282	0.024758	-49.04662	0.0000
AR(3)	0.804555	0.020288	39.65703	0.0000
AR(4)	-1.257104	0.020396	-61.63585	0.0000
AR(5)	1.184911	0.027069	43.77372	0.0000
AR(6)	-0.549937	0.020739	-26.51642	0.0000
MA(1)	0.115114	0.010145	11.34708	0.0000
MA(2)	-0.046410	0.009030	-5.139412	0.0000
MA(3)	-0.173108	0.010570	-16.37678	0.0000
MA(4)	-0.887668	0.009971	-89.02887	0.0000
Variance Equation				
C	9.85E-10	2.32E-10	4.248112	0.0000
RESID(-1)^2	0.265391	0.032790	8.093698	0.0000
GARCH(-1)	0.703556	0.031159	22.57985	0.0000
GED PARAMETER	1.855356	0.103766	17.88011	0.0000
R-squared	0.974126	Mean dependent var	1.88E-06	
Adjusted R-squared	0.974008	S.D. dependent var	0.000873	
S.E. of regression	0.000141	Akaike info criterion	-15.11198	
Sum squared resid	3.91E-05	Schwarz criterion	-15.07250	
Log likelihood	14997.53	Hannan-Quinn criter.	-15.09748	
Durbin-Watson stat	1.975192			

Heteroskedasticity Test: ARCH

F-statistic	0.159476	Prob. F(1,1980)	0.6897
Obs*R-squared	0.159624	Prob. Chi-Square(1)	0.6895

(2) Modelo ARIMA-GARCH para a componente de detalhe $\left(\tilde{f}_{D_2}(t)\right)_{t=1}^{1994}$, onde o modelo ARIMA é dado por:

$$\begin{aligned} \hat{f}_{D_2}(t) = & \left[-\tilde{f}_{D_2}(t-1) \times 2.265951 - \tilde{f}_{D_2}(t-2) \times 3.369826 - \tilde{f}_{D_2}(t-3) \times 3.715283 \right. \\ & - \tilde{f}_{D_2}(t-4) \times 3.322687 - \tilde{f}_{D_2}(t-5) \times 2.308898 - \tilde{f}_{D_2}(t-6) \times 1.305184 \\ & \left. - \tilde{f}_{D_2}(t-7) \times 0.537462 + \tilde{f}_{D_2}(t-8) \times 0.133859 \right] + \left[-\hat{e}_{D_2}(t- \right. \end{aligned}$$

$$1) \times 1.494498 + \hat{e}_{D_2}(t-3) \times 0.444167 + \hat{e}_{D_2}(t-4) \times 0.691475 - \hat{e}_{D_2}(t-5) \times 0.611763] + \hat{\sigma}_{D_2,t}^2 \times 0.$$

e o modelo GARCH (1,1), com distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{A_2,t}^2 = 0.093152 - 0.000782 \times (\hat{e}_{A_2}(t-1))^2 - 0.995189 \times \hat{\sigma}_{A_2,t-1}^2.$$

Dependent Variable: D2_F13X

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	-2.265951	0.015880	-142.6898	0.0000
AR(2)	-3.369826	0.055618	-60.58926	0.0000
AR(3)	-3.715283	0.071960	-51.62948	0.0000
AR(4)	-3.322687	0.088145	-37.69556	0.0000
AR(5)	-2.308898	0.080070	-28.83614	0.0000
AR(6)	-1.305184	0.054803	-23.81603	0.0000
AR(7)	-0.537462	0.030379	-17.69192	0.0000
AR(8)	-0.133859	0.011288	-11.85876	0.0000
MA(1)	-1.494498	0.014246	-104.9034	0.0000
MA(3)	0.444167	0.060958	7.286411	0.0000
MA(4)	0.691475	0.078941	8.759400	0.0000
MA(5)	-0.611763	0.031680	-19.31068	0.0000
Variance Equation				
C	0.093152	9774.624	18.34497	0.0000
RESID(-1)^2	-0.000782	7.11E-05	-10.99749	0.0000
GARCH(-1)	-0.995189	0.001264	-787.4379	0.0000
GED PARAMETER	0.692167	0.019459	35.57114	0.0000
R-squared	0.982283	Mean dependent var	-1.664819	
Adjusted R-squared	0.982185	S.D. dependent var	2714.408	
S.E. of regression	362.3041	Akaike info criterion	13.83753	
Sum squared resid	2.59E+08	Schwarz criterion	13.88264	
Log likelihood	-13710.83	Hannan-Quinn criter.	13.85410	
Durbin-Watson stat	2.334959			

(3) Modelo ARIMA-GARCH para a componente de detalhe $\left(\tilde{f}_{D_3}(t)\right)_{t=1}^{1994}$, onde o modelo ARIMA é dado por:

$$\hat{f}_{D_3}(t) = [-\tilde{f}_{D_3}(t-2) \times 1.001961 - \tilde{f}_{D_3}(t-4) \times 0.467598 - \tilde{f}_{D_3}(t-6) \times 0.468774 + \tilde{f}_{D_3}(t-8) \times 0.000529 - \tilde{f}_{D_3}(t-16) \times 0.000392 + \tilde{f}_{D_3}(t-$$

$$20) \times 0.000199] + [\hat{e}_{D_3}(t-1) \times 0.997975 - \hat{e}_{D_3}(t-10) \times 0.000303 - \hat{e}_{D_3}(t-12) \times 0.000251]$$

e o modelo GARCH (1,1), com ordem assimétrica igual a 1 e distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{D_2,t}^2 = \left(\hat{e}_{A_2}(t-1)\right)^2 \times 0.008614 + \hat{\sigma}_{D_2,t-1}^2 \times 0.991386.$$

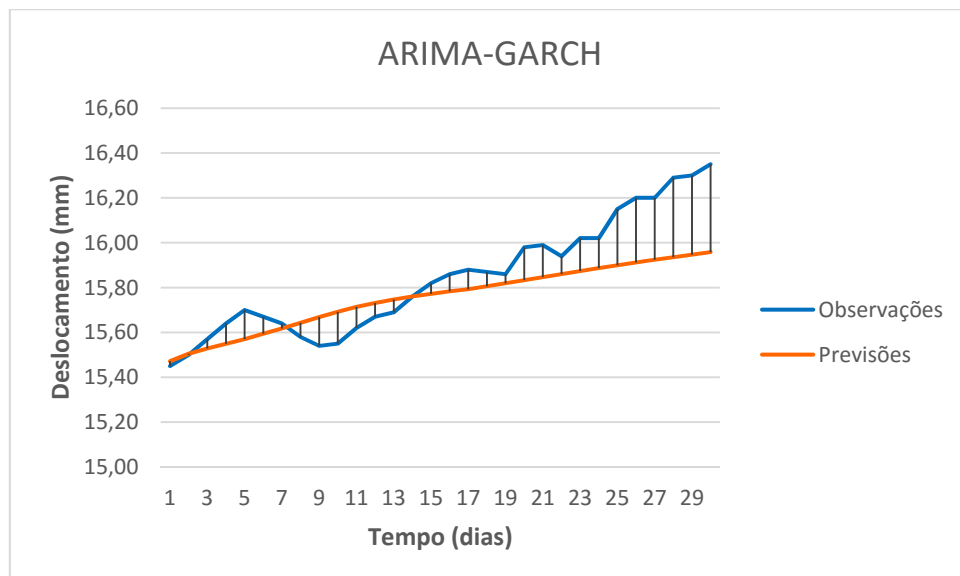
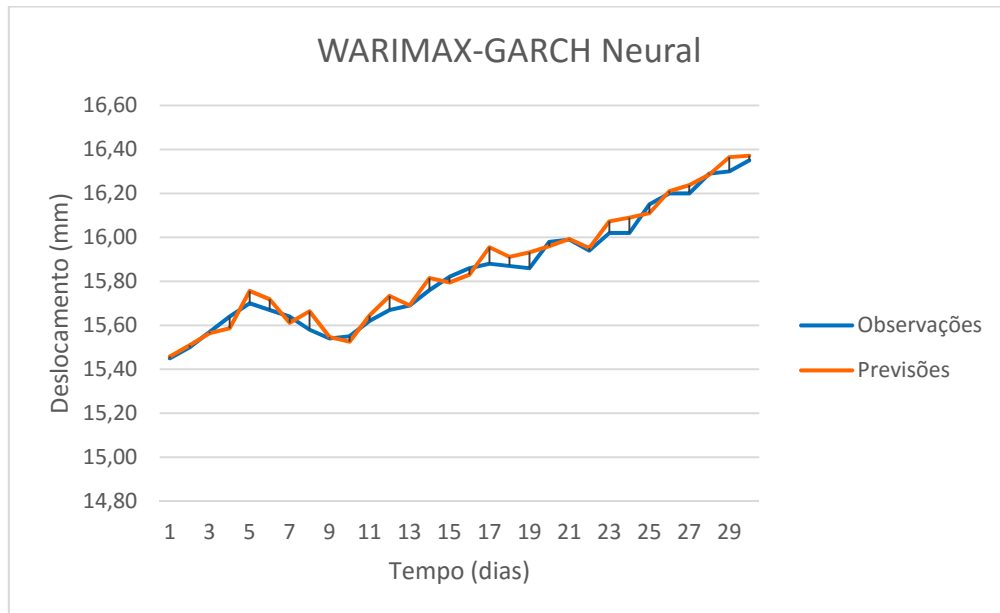
Dependent Variable: D3_F13X

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(2)	-1.001961	9.04E-06	-110829.9	0.0000
AR(4)	-0.467598	0.000175	-2668.816	0.0000
AR(6)	-0.468774	0.000176	-2664.538	0.0000
AR(8)	-0.000529	7.34E-06	-72.11968	0.0000
AR(16)	-0.000392	3.44E-06	-113.8552	0.0000
AR(20)	0.000199	3.00E-06	66.32334	0.0000
MA(1)	0.997975	3.58E-06	278795.6	0.0000
MA(10)	-0.000303	4.05E-06	-74.91943	0.0000
MA(12)	-0.000251	2.45E-06	-102.0972	0.0000
Variance Equation				
RESID(-1)^2	0.008614	0.001989	4.330596	0.0000
GARCH(-1)	0.991386	0.001989	498.4359	0.0000
GED PARAMETER	0.145528	0.002103	69.19668	0.0000
R-squared	0.653902	Mean dependent var		0.000000
Adjusted R-squared	0.652594	S.D. dependent var		0.059407
S.E. of regression	0.035015	Akaike info criterion		-9.399437
Sum squared resid	2.920480	Schwarz criterion		-9.370441
Log likelihood	11253.73	Hannan-Quinn criter.		-9.388886
Durbin-Watson stat	1.999633			

Apêndice III E: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX-GARCH Neural (F13x).

Na Figura 61, é feita a comparação da evolução temporal das previsões fora da amostra para WARIMAX-GARCH Neural, RNA e ARIMA-GARCH em relação às observações de fora da amostra



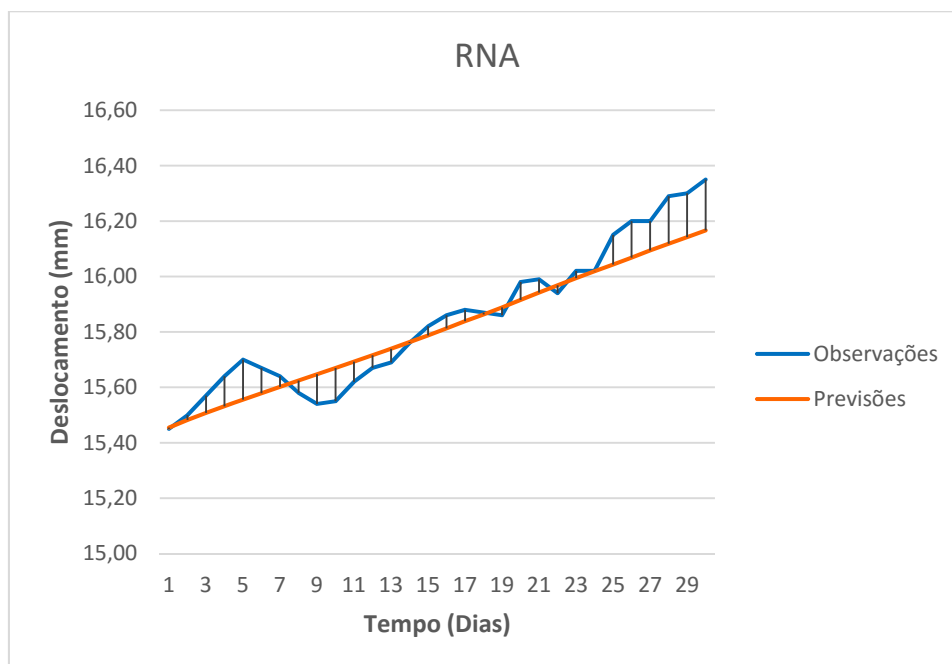


Figura 61 – Evolução temporal do erro APE para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA

Dependent Variable: F13X

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
A2_F13XF(-1)	3.424953	0.000221	15517.73	0.0000
A2_F13XF(-2)	-4.808460	0.000387	-12436.40	0.0000
D2_F13XF(-1)	3.331650	0.000113	29561.56	0.0000
D3_F13XF(-1)	-0.948146	0.000319	-2967.847	0.0000
AR(2)	0.295259	0.014685	20.10585	0.0000
AR(3)	-0.696244	0.014382	-48.40921	0.0000
MA(2)	-0.908864	0.002548	-356.7333	0.0000
MA(3)	0.086151	0.001697	50.77687	0.0000
Variance Equation				
C	0.000169	2.88E-05	5.881009	0.0000
RESID(-1)^2	0.291072	0.029920	9.728216	0.0000
GARCH(-1)	0.550262	0.038262	14.38161	0.0000
GED PARAMETER	1.507477	0.049339	30.55363	0.0000
R-squared	0.999384	Mean dependent var	14.89727	
Adjusted R-squared	0.999382	S.D. dependent var	1.440872	
S.E. of regression	0.035817	Akaike info criterion	-4.329916	
Sum squared resid	2.537562	Schwarz criterion	-4.296116	
Log likelihood	4311.607	Hannan-Quinn criter.	-4.317501	
Durbin-Watson stat	2.092066			

Heteroskedasticity Test: ARCH

F-statistic	0.001407	Prob. F(1,1983)	0.9701
Obs*R-squared	0.001408	Prob. Chi-Square(1)	0.9701

Apêndice IV A: Teste Dikey-Fuller Aumentado (F13y).

Null Hypothesis: F13Y has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.018272	0.0014
Test critical values: 1% level	-3.433416	
5% level	-2.862781	
10% level	-2.567477	

Null Hypothesis: D(F13Y) has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-15.37292	0.0000
Test critical values: 1% level	-3.433400	
5% level	-2.862773	
10% level	-2.567473	

Apêndice IV B: Configurações da RNA (série temporal F13y).

Tabela 10 – Configurações das RNA-MLPs para as componentes *wavelet*.

Parâmetros da arquitetura neural	Componente de aproximação nível 2
Transformação	premnmx
Tamanho de janela de entrada	5
Número de camadas escondidas	1
Número de neurônio na camada escondida	2
Função de ativação da camada oculta	Tangente hiperbólica
Função de ativação da camada de saída	Linear
Algoritmo de treinamento	Levenberg-Marquardt (Millar, 2011)

Apêndice IV C: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (série temporal F13y).

O modelo ARIMA (1,0,5) é dado por:

$$\hat{f}(t) = [\tilde{f}(t-1) \times 1.834345 - \tilde{f}(t-2) \times 0.789061 - \tilde{f}(t-5) \times 0.045323] + [-\hat{e}(t-1) \times 0.891779 + \hat{e}(t-2) \times 0.040703 + \hat{e}(t-3) \times 0.063133]$$

Por sua vez, o modelo GARCH (2,1) é dado por:

$$\hat{\sigma}_t^2 = 1.31 \times 10^{-5} + (\hat{e}(t-1))^2 \times 0.223499 - (\hat{e}(t-2))^2 \times 0.161167 + \hat{\sigma}_{A_2t-1}^2 \times 0.921920$$

Dependent Variable: F13Y

Method: ML - ARCH (Marquardt) - Student's t distribution

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	1.834345	0.001734	1057.873	0.0000
AR(2)	-0.789061	0.002299	-343.1589	0.0000
AR(5)	-0.045323	9.34E-05	-485.0580	0.0000
MA(1)	-0.891779	0.025372	-35.14787	0.0000
MA(2)	0.040703	0.031391	1.296656	0.1947
MA(3)	-0.063133	0.021306	-2.963202	0.0030
Variance Equation				
C	1.31E-05	4.57E-06	2.870364	0.0041
RESID(-1)^2	0.223499	0.040051	5.580417	0.0000
RESID(-2)^2	-0.161167	0.040535	-3.976016	0.0001
GARCH(-1)	0.921920	0.015862	58.12056	0.0000
T-DIST. DOF	6.198014	0.750906	8.254049	0.0000
R-squared	0.997004	Mean dependent var	-2.571308	
Adjusted R-squared	0.996997	S.D. dependent var	0.526725	
S.E. of regression	0.028866	Akaike info criterion	-4.514400	
Sum squared resid	1.651519	Schwarz criterion	-4.483442	
Log likelihood	4498.314	Hannan-Quinn criter.	-4.503030	
Durbin-Watson stat	1.796225			

Heteroskedasticity Test: ARCH

F-statistic	1.863183	Prob. F(1,1985)	0.1724
Obs*R-squared	1.863311	Prob. Chi-Square(1)	0.1722

Apêndice IV D: Modelo WARIMAX-GARCH Neural e seus resultados estatísticos para validação (série temporal F13y).

(1) Modelo ARIMA-GARCH para a componente de aproximação $\left(\tilde{f}_{A_2}(t)\right)_{t=1}^{1993}$,

onde o modelo ARIMA é dado por:

$$\nabla \left(\hat{f}_{A_2}(t) \right) = \left[\tilde{f}_{A_2}(t-1) \times 2.85494 - \tilde{f}_{A_2}(t-2) \times 2.676184 + \tilde{f}_{A_2}(t-4) \times 1.231500 - \tilde{f}_{A_2}(t-6) \times 0.615992 + \tilde{f}_{A_2}(t-8) \times 0.201866 - \tilde{f}_{A_2}(t-12) \times 0.026503 \right] + \left[0 \times \hat{\sigma}_{A_2,t}^2 \right]$$

e o modelo GARCH (1,1), com distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{A_2,t}^2 = 2.77 \times 10^{-11} + 1.140590 \times \left(\hat{e}_{A_2}(t-1) \right)^2 + 0.236298 \times \hat{\sigma}_{A_2,t-1}^2.$$

Dependent Variable: D(A2_F13Y)

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	2.854947	0.004624	617.4711	0.0000
AR(2)	-2.676184	0.007485	-357.5292	0.0000
AR(4)	1.231500	0.005916	208.1645	0.0000
AR(6)	-0.615992	0.003857	-159.6983	0.0000
AR(8)	0.201866	0.001493	135.2196	0.0000
AR(12)	-0.026503	0.000440	-60.18969	0.0000
Variance Equation				
C	2.77E-11	9.23E-12	3.001194	0.0027
RESID(-1)^2	1.140590	0.086262	13.22236	0.0000
GARCH(-1)	0.236298	0.015651	15.09834	0.0000
GED PARAMETER	1.471267	0.077946	18.87547	0.0000
R-squared	0.998932	Mean dependent var	-1.35E-05	
Adjusted R-squared	0.998927	S.D. dependent var	0.016658	
S.E. of regression	0.000546	Akaike info criterion	-14.13104	
Sum squared resid	0.000478	Schwarz criterion	-14.09101	
Log likelihood	11422.81	Hannan-Quinn criter.	-14.11618	
Durbin-Watson stat	2.060502			

Heteroskedasticity Test: ARCH

F-statistic	1.100911	Prob. F(1,1970)	0.2942
Obs*R-squared	1.101414	Prob. Chi-Square(1)	0.2940

(2) Modelo ARIMA para a componente de detalhe $\left(\tilde{f}_{D_2}(t)\right)_{t=1}^{1993}$ é dado por:

$$\begin{aligned} \hat{f}_{D_2}(t) = & \left[-\tilde{f}_{D_2}(t-1) \times 1.646784 - \tilde{f}_{D_2}(t-2) \times 1.682723 - \tilde{f}_{D_2}(t-3) \times 0.840341 - \tilde{f}_{D_2}(t-8) \times 0.091690 \right] + \\ & \left[-\hat{e}_{D_2}(t-1) \times 1.334413 + \hat{e}_{D_2}(t-3) \times 0.332709 + \hat{e}_{D_2}(t-4) \times 0.225037 - \hat{e}_{D_2}(t-5) \times 0.077859 \right] + \\ & \hat{\sigma}_{D_2,t}^2 \times 0. \end{aligned}$$

e o modelo GARCH (1,1) com distribuição do erro generalizada (GED) é dado por:

$$\hat{\sigma}_{D_2,t}^2 = 2.76 \times 10^{-7} - \left(\hat{e}_{A_2}(t-1) \right)^2 \times 0.348814 - \hat{\sigma}_{D_2,t-1}^2 \times 0.615269.$$

Dependent Variable: D2_F13Y

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	-1.646784	0.066673	-24.69938	0.0000
AR(2)	-1.682723	0.199410	-8.438507	0.0000
AR(3)	-0.840341	0.359420	-2.338051	0.0194
AR(8)	-0.091690	0.042729	-2.145844	0.0319
MA(1)	-1.334413	0.006103	-218.6636	0.0000
MA(3)	0.332709	0.001639	202.9896	0.0000
MA(4)	0.225037	0.004921	45.73083	0.0000
MA(5)	-0.077859	0.000544	-143.0383	0.0000
Variance Equation				
C	2.76E-07	5.88E-08	4.686178	0.0000
RESID(-1)^2	0.348814	0.039345	8.865414	0.0000
GARCH(-1)	0.615269	0.035309	17.42546	0.0000
GED PARAMETER	1.848454	0.081848	22.58407	0.0000
R-squared	0.972771	Mean dependent var	-1.42E-06	
Adjusted R-squared	0.972619	S.D. dependent var	0.011118	
S.E. of regression	0.001840	Akaike info criterion	-10.00807	
Sum squared resid	0.006678	Schwarz criterion	-9.962986	
Log likelihood	9949.010	Hannan-Quinn criter.	-9.991510	
Durbin-Watson stat	2.042067			

Heteroskedasticity Test: ARCH

F-statistic	4.945032	Prob. F(1,1982)	0.0263
Obs*R-squared	4.937703	Prob. Chi-Square(1)	0.0263

(3) Modelo ARIMA-GARCH para a componente de detalhe $\left(\tilde{f}_{D_3}(t)\right)_{t=1}^{1993}$, onde o modelo ARIMA é dado por:

$$\begin{aligned} \hat{f}_{D_3}(t) = & [\tilde{f}_{D_3}(t-1) \times 0.880721 - \tilde{f}_{D_3}(t-2) \times 2.147480 - \tilde{f}_{D_3}(t-3) \times 0.269246 + \tilde{f}_{D_3}(t-4) \times 0.895915 - \tilde{f}_{D_3}(t-5) \times 1.410807 - \tilde{f}_{D_3}(t-7) \times 0.928008 + \tilde{f}_{D_3}(t-8) \times 0.527949 + \tilde{f}_{D_3}(t-10) \times 0.353120] + \\ & [\hat{e}_{D_3}(t-1) \times 0.735844 - \hat{e}_{D_3}(t-2) \times 1.718650 - \hat{e}_{D_3}(t-3) \times 1.961380 + \hat{e}_{D_3}(t-4) \times 0.524908 + \hat{e}_{D_3}(t-5) \times 1.065598 + \hat{e}_{D_3}(t-7) \times 0.451835 + \hat{e}_{D_3}(t-8) \times 0.774244 - \hat{e}_{D_3}(t-9) \times 0.327392 - \hat{e}_{D_3}(t-10) \times 0.544504] \end{aligned}$$

e o modelo GARCH (1,1), com ordem assimétrica igual a 1 e distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{D_2,t}^2 = 4.19^{-9} - \left(\hat{e}_{A_2}(t-1)\right)^2 \times 0.182185 - \hat{\sigma}_{D_2,t-1}^2 \times 0.796036.$$

Dependent Variable: D2_F13Y

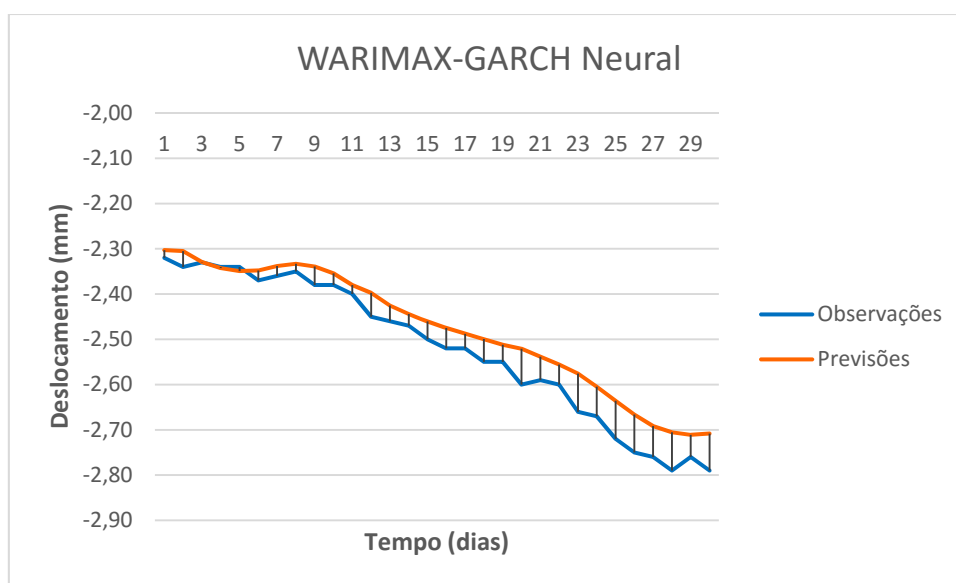
Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(2)	-2.024690	0.016886	-119.9023	0.0000
AR(4)	-2.394922	0.027243	-87.90997	0.0000
AR(6)	-1.727336	0.027259	-63.36860	0.0000
AR(8)	-0.652989	0.016614	-39.30332	0.0000
AR(16)	-0.060434	0.005289	-11.42718	0.0000
AR(20)	0.026914	0.005090	5.287935	0.0000
MA(1)	1.143165	0.004634	246.7007	0.0000
MA(10)	0.306300	0.024585	12.45890	0.0000
MA(12)	-0.161947	0.020050	-8.077101	0.0000
Variance Equation				
RESID(-1)^2	0.230841	0.010927	21.12573	0.0000
GARCH(-1)	0.769159	0.010927	70.39072	0.0000
GED PARAMETER	2.117021	0.100154	21.13762	0.0000

R-squared	0.963835	Mean dependent var	3.88E-06
Adjusted R-squared	0.963670	S.D. dependent var	0.013676
S.E. of regression	0.002607	Akaike info criterion	-9.609730
Sum squared resid	0.013338	Schwarz criterion	-9.575747
Log likelihood	9491.999	Hannan-Quinn criter.	-9.597244
Durbin-Watson stat	1.951566		

Apêndice IV E: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX-GARCH Neural (F13y).

Na Figura 62, podemos comparar a evolução temporal das previsões fora da amostra para WARIMAX-GARCH Neural e ARIMA-GARCH em relação às observações de fora da amostra



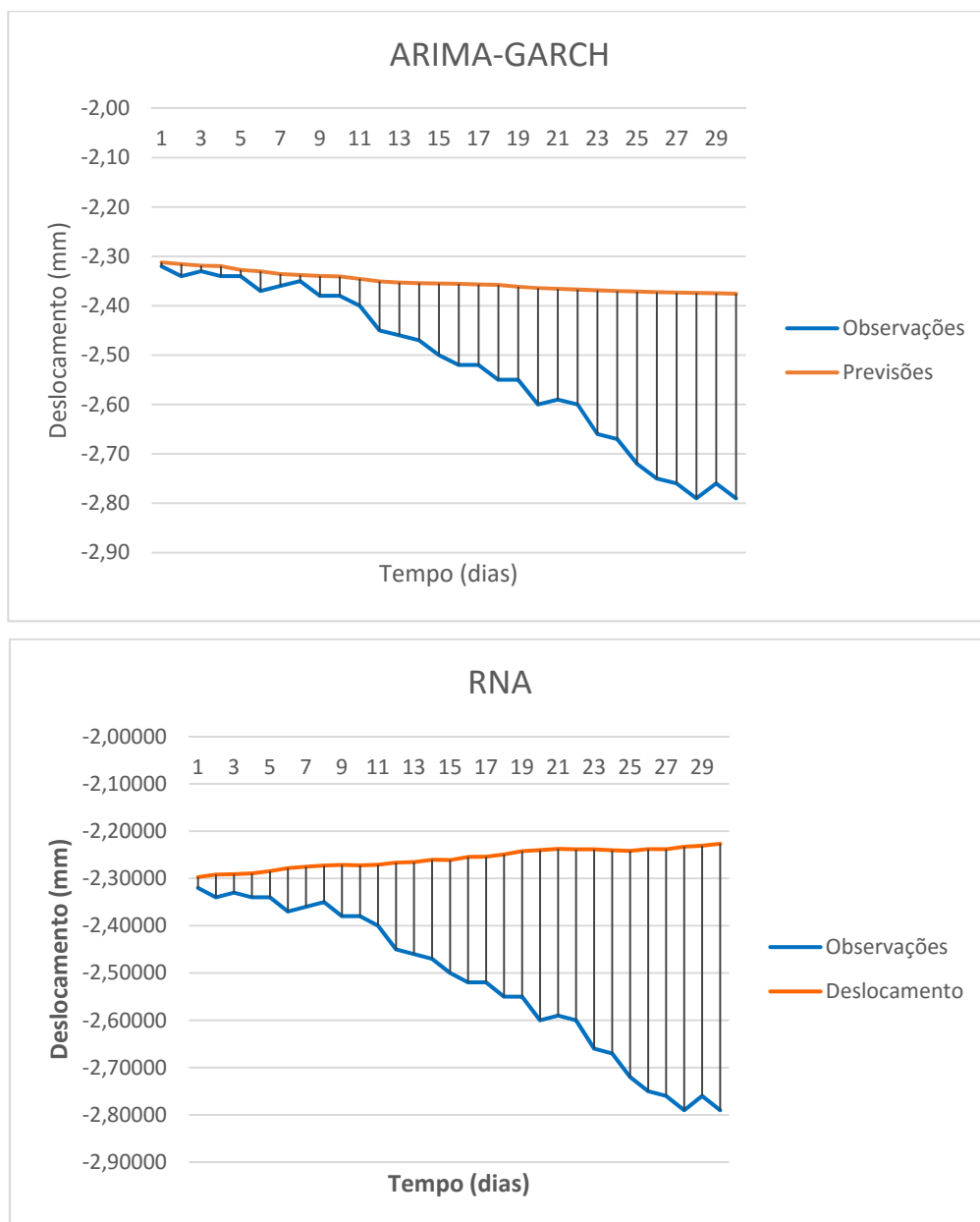


Figura 62 – Evolução temporal do erro APE para WARIMAX-GARCH Neural, ARIMA-GARCH e RNA

Dependent Variable: D(F13Y,2)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
A2_F13YC(-1)	1.602457	0.026379	60.74745	0.0000
A2_F13YC(-2)	-3.926994	0.077028	-50.98163	0.0000
D2_F13YC(-1)	3.046077	0.076717	39.70527	0.0000
D3_F13YC(-1)	-0.721539	0.026036	-27.71348	0.0000
AR(1)	-0.752077	0.023928	-31.43092	0.0000
AR(2)	-0.755698	0.027374	-27.60648	0.0000
AR(3)	-0.688668	0.028260	-24.36923	0.0000

AR(4)	-0.498271	0.027624	-18.03742	0.0000
AR(5)	-0.247473	0.023933	-10.34023	0.0000
MA(1)	-0.969601	0.008996	-107.7774	0.0000

Variance Equation				
C	1.21E-05	3.34E-06	3.611618	0.0003
RESID(-1)^2	0.115281	0.017885	6.445584	0.0000
GARCH(-1)	0.853887	0.021346	40.00241	0.0000

T-DIST. DOF	8.144423	1.203494	6.767312	0.0000
-------------	----------	----------	----------	--------

R-squared	0.760704	Mean dependent var	-2.52E-05
Adjusted R-squared	0.759613	S.D. dependent var	0.039618
S.E. of regression	0.019424	Akaike info criterion	-5.269184
Sum squared resid	0.744809	Schwarz criterion	-5.229718
Log likelihood	5241.030	Hannan-Quinn criter.	-5.254687
Durbin-Watson stat	1.986453		

Heteroskedasticity Test: ARCH

F-statistic	1.875769	Prob. F(1,1981)	0.1710
Obs*R-squared	1.875887	Prob. Chi-Square(1)	0.1708

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 0.001	0.001	0.0011	
		2 -0.002	-0.002	0.0135	
		3 0.001	0.001	0.0176	
		4 -0.001	-0.001	0.0191	
		5 0.001	0.001	0.0234	
		6 0.000	0.000	0.0236	
		7 0.001	0.001	0.0259	
		8 0.004	0.004	0.0763	
		9 0.001	0.001	0.0776	
		10 -0.003	-0.003	0.0997	0.752
		11 0.002	0.002	0.1086	0.947
		12 -0.000	-0.000	0.1087	0.991
		13 0.001	0.001	0.1136	0.998
		14 -0.000	-0.000	0.1137	1.000
		15 0.001	0.001	0.1174	1.000
		16 -0.000	-0.000	0.1175	1.000
		17 0.001	0.001	0.1193	1.000
		18 -0.000	-0.000	0.1195	1.000
		19 0.001	0.001	0.1211	1.000
		20 -0.002	-0.002	0.1344	1.000
		21 0.002	0.002	0.1403	1.000
		22 0.000	0.000	0.1404	1.000
		23 0.001	0.001	0.1431	1.000
		24 0.000	0.000	0.1431	1.000
		25 0.001	0.001	0.1467	1.000
		26 0.000	0.000	0.1467	1.000
		27 0.000	0.000	0.1467	1.000
		28 0.000	0.000	0.1467	1.000
		29 0.001	0.001	0.1488	1.000
		30 0.000	0.000	0.1489	1.000
		31 0.000	0.000	0.1491	1.000
		32 -0.000	-0.000	0.1491	1.000
		33 0.000	0.000	0.1491	1.000
		34 0.000	0.000	0.1491	1.000
		35 -0.000	-0.000	0.1491	1.000
		36 0.000	0.000	0.1492	1.000

FAC e FACP dos resíduos
ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 0.031	0.031	1.8795	0.170
		2 0.001	0.000	1.8821	0.390
		3 0.005	0.005	1.9396	0.585
		4 0.008	0.008	2.0762	0.722
		5 -0.008	-0.009	2.2197	0.818
		6 -0.018	-0.017	2.8460	0.828
		7 -0.027	-0.026	4.2889	0.746
		8 0.035	0.036	6.6800	0.572
		9 -0.004	-0.006	6.7145	0.667
		10 -0.026	-0.025	8.0763	0.621
		11 -0.005	-0.003	8.1217	0.702
		12 0.012	0.011	8.4139	0.752
		13 -0.004	-0.004	8.4389	0.814
		14 -0.036	-0.035	10.969	0.688
		15 0.019	0.023	11.720	0.700
		16 0.032	0.028	13.749	0.617
		17 0.017	0.015	14.340	0.643
		18 0.014	0.016	14.757	0.679
		19 -0.016	-0.018	15.271	0.705
		20 0.000	-0.002	15.271	0.761
		21 -0.001	-0.002	15.275	0.809
		22 0.003	0.008	15.289	0.850
		23 0.002	0.002	15.297	0.884
		24 -0.024	-0.027	16.435	0.872
		25 -0.025	-0.023	17.684	0.856
		26 -0.027	-0.025	19.111	0.832
		27 -0.011	-0.007	19.351	0.857
		28 0.001	0.002	19.353	0.887
		29 0.004	0.005	19.386	0.911
		30 -0.037	-0.038	22.211	0.846
		31 0.017	0.017	22.768	0.857
		32 -0.006	-0.007	22.835	0.883
		33 -0.007	-0.009	22.936	0.905
		34 0.021	0.021	23.799	0.904
		35 -0.004	-0.007	23.835	0.923
		36 -0.002	-0.003	23.845	0.940

FAC e FACP dos resíduos
quadráticos padronizados

Apêndice V A: Teste Dikey-Fuller Aumentado (série temporal F19x).

Null Hypothesis: F19X has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.732858	0.4145
Test critical values: 1% level	-3.433400	
5% level	-2.862773	
10% level	-2.567473	

Null Hypothesis: D(F19X) has a unit root

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-22.71591	0.0000
Test critical values: 1% level	-3.433400	
5% level	-2.862773	
10% level	-2.567473	

Apêndice V B: Configurações da RNA (série temporal F19x).

Tabela 11 – Configurações das RNA-MLPs para as componentes *wavelet*.

Parâmetros da arquitetura neural	Componente de aproximação nível 2
Transformação	premnmx
Tamanho de janela de entrada	6
Número de camadas escondidas	1
Número de neurônio na camada escondida	4
Função de ativação da camada oculta	Tangente hiperbólica
Função de ativação da camada de saída	Linear
Algoritmo de treinamento	Levenberg-Marquardt (Millar, 2011)

Apêndice V C: Modelo ARIMA-GARCH e seus resultados estatísticos para validação (série temporal F19x).

O modelo ARIMA (7,1,0) é dado por:

$$\nabla(\hat{f}(t)) = [\tilde{f}(t-1) \times 0.396301 + \tilde{f}(t-2) \times 0.084378 + \tilde{f}(t-7) \times 0.075636]$$

onde ∇ denota o operador diferença de primeira ordem não sazonal. Por sua vez, o modelo GARCH (1,1) é dado por:

$$\hat{\sigma}_t^2 = 0.000289 + (\hat{e}(t-1))^2 \times 0.084274 + \hat{\sigma}_{A_2t-1}^2 \times 0.846311.$$

Dependent Variable: D(F19X)

Method: ML - ARCH (Marquardt) - Student's t distribution

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.396301	0.022109	17.92481	0.0000
AR(2)	0.084378	0.022916	3.682090	0.0002
AR(7)	0.075636	0.020149	3.753857	0.0002
Variance Equation				
C	0.000289	9.34E-05	3.094604	0.0020
RESID(-1)^2	0.084274	0.018821	4.477563	0.0000
GARCH(-1)	0.846311	0.033647	25.15261	0.0000
T-DIST. DOF	7.876723	1.268554	6.209216	0.0000
R-squared	0.160638	Mean dependent var		-0.000358
Adjusted R-squared	0.159791	S.D. dependent var		0.069538
S.E. of regression	0.063740	Akaike info criterion		-2.736359
Sum squared resid	8.052547	Schwarz criterion		-2.716634
Log likelihood	2722.836	Hannan-Quinn criter.		-2.729114
Durbin-Watson stat	2.093550			

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	0.024	-0.024	1.1737		
2	-0.008	-0.009	1.3072		
3	-0.022	-0.022	2.2443		
4	-0.006	-0.007	2.3059		
5	0.006	0.006	2.3825		
6	0.006	0.005	2.4450		
7	-0.022	-0.022	3.3773	0.066	
8	-0.009	-0.009	3.5238	0.172	
9	-0.017	-0.018	4.1192	0.249	
10	0.047	0.045	8.4855	0.075	
11	-0.021	-0.020	9.3973	0.094	
12	0.036	0.036	12.046	0.061	
13	-0.005	-0.002	12.094	0.097	
14	0.009	0.009	12.246	0.141	
15	0.008	0.009	12.389	0.192	
16	0.009	0.009	12.542	0.250	
17	0.005	0.008	12.597	0.320	
18	0.000	0.000	12.597	0.399	
19	0.054	0.057	18.408	0.143	
20	-0.002	-0.001	18.416	0.189	
21	0.053	0.058	23.964	0.066	
22	0.031	0.033	25.925	0.055	
23	0.008	0.015	28.060	0.073	
24	0.025	0.027	27.287	0.074	
25	0.005	0.008	27.330	0.097	
26	-0.013	-0.010	27.672	0.117	
27	0.034	0.034	29.937	0.093	
28	0.003	0.009	29.954	0.120	
29	0.028	0.026	31.517	0.111	
30	0.034	0.043	33.875	0.087	
31	0.024	0.021	35.005	0.088	
32	0.030	0.035	36.787	0.078	
33	-0.009	-0.008	36.939	0.096	
34	0.036	0.036	39.620	0.071	
35	0.011	0.013	39.870	0.086	
36	-0.012	-0.010	40.139	0.102	

FAC e FACP dos resíduos
ordinários padronizados

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
1	0.057	0.057	6.5604	0.010	
2	-0.012	-0.015	6.8300	0.033	
3	-0.020	-0.018	7.6129	0.055	
4	-0.031	-0.029	9.4929	0.050	
5	-0.027	-0.024	10.981	0.052	
6	-0.009	-0.007	11.127	0.085	
7	-0.016	-0.017	11.635	0.113	
8	-0.025	-0.025	12.839	0.118	
9	-0.027	-0.026	14.244	0.114	
10	-0.031	-0.031	16.181	0.095	
11	0.002	0.003	16.191	0.134	
12	0.001	-0.003	16.195	0.182	
13	-0.014	-0.018	16.590	0.219	
14	0.006	0.004	16.655	0.275	
15	0.059	0.056	23.651	0.071	
16	0.037	0.029	26.362	0.049	
17	0.007	0.002	26.465	0.066	
18	-0.006	-0.006	26.537	0.088	
19	0.010	0.015	26.750	0.111	
20	-0.029	-0.026	28.383	0.101	
21	-0.006	-0.000	28.446	0.128	
22	-0.026	-0.025	29.774	0.124	
23	0.005	0.010	29.820	0.155	
24	0.029	0.031	31.504	0.140	
25	0.009	0.009	31.668	0.168	
26	0.029	0.029	33.389	0.151	
27	-0.027	-0.031	34.873	0.142	
28	-0.007	0.000	34.963	0.171	
29	0.055	0.058	41.108	0.067	
30	0.031	0.020	42.986	0.059	
31	-0.010	-0.016	43.202	0.071	
32	0.011	0.014	43.452	0.085	
33	-0.009	-0.004	43.610	0.102	
34	0.016	0.024	44.143	0.114	
35	-0.003	-0.002	44.167	0.138	
36	-0.027	-0.023	45.683	0.129	

FAC e FACP dos resíduos
quadráticos padronizados

Heteroskedasticity Test: ARCH

F-statistic	0.902658	Prob. F(1,2484)	0.3422
Obs*R-squared	0.903057	Prob. Chi-Square(1)	0.3420

Apêndice V D: Modelo WARIMAX-GARCH Neural e seus resultados estatísticos para validação (série temporal F19x).

(1) Modelo ARIMA-GARCH para a componente de aproximação $\left(\tilde{f}_{A_2}(t)\right)_{t=1}^{1993}$,

onde o modelo ARIMA é dado por:

$$\nabla\left(\hat{f}_{A_2}(t)\right)=\left[\tilde{f}_{A_2}(t-1) \times 1.659544-\tilde{f}_{A_2}(t-2) \times 0.542116-\tilde{f}_{A_2}(t-4) \times 0.888710+\tilde{f}_{A_2}(t-5) \times 1.156196-\tilde{f}_{A_2}(t-7) \times 0.361865+\tilde{f}_{A_2}(t-9) \times 0.171455\right]+\left[\hat{e}_{A_2}(t-1) \times 2.254992+\hat{e}_{A_2}(t-2) \times 2.312326-\hat{e}_{A_2}(t-3) \times 1.766190-\hat{e}_{A_2}(t-5) \times 1.580044-\hat{e}_{A_2}(t-6) \times 1.560172-\right.$$

$$\hat{e}_{A_2}(t-7) \times 0.946965 + \hat{e}_{A_2}(t-9) \times 0.588999 + \hat{e}_{A_2}(t-10) \times 0.400107 + 0 \times \hat{\sigma}_{A_2,t}^2 \Big]$$

e o modelo GARCH (1,1) e distribuição do erro generalizada (GED), é dado por:

$$\hat{\sigma}_{A_2,t}^2 = 1.32 \times 10^{-9} + 0.291737 \times \left(\hat{e}_{A_2}(t-1) \right)^2 + 0.569394 \times \hat{\sigma}_{A_2,t-1}^2.$$

Dependent Variable: D(LOG(A2_F19X))

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	1.659544	0.103883	15.97510	0.0000
AR(2)	-0.542116	0.145056	-3.737282	0.0002
AR(4)	-0.888710	0.094972	-9.357602	0.0000
AR(5)	1.156196	0.094574	12.22529	0.0000
AR(7)	-0.361865	0.092354	-3.918248	0.0001
AR(9)	0.171455	0.027440	6.248304	0.0000
MA(1)	2.254992	0.103410	21.80632	0.0000
MA(2)	2.312326	0.282388	8.188460	0.0000
MA(3)	1.766190	0.381984	4.623734	0.0000
MA(5)	-1.580044	0.252502	-6.257553	0.0000
MA(6)	-1.560172	0.214299	-7.280343	0.0000
MA(7)	-0.946965	0.315067	-3.005594	0.0027
MA(9)	0.588999	0.225624	2.610530	0.0090
MA(10)	0.400107	0.075732	5.283193	0.0000
Variance Equation				
C	1.32E-09	2.44E-10	5.413565	0.0000
RESID(-1)^2	0.291737	0.039266	7.429708	0.0000
GARCH(-1)	0.569394	0.048579	11.72098	0.0000
GED PARAMETER	1.868838	0.091774	20.36341	0.0000
R-squared	0.999160	Mean dependent var	-1.46E-05	
Adjusted R-squared	0.999152	S.D. dependent var	0.003171	
S.E. of regression	9.23E-05	Akaike info criterion	-15.86336	
Sum squared resid	1.67E-05	Schwarz criterion	-15.79850	
Log likelihood	15751.52	Hannan-Quinn criter.	-15.83954	
Durbin-Watson stat	1.837643			

(2) Modelo ARIMA para a componente de detalhe $\left(\tilde{f}_{D_2}(t) \right)_{t=1}^{1993}$ é dado por:

$$\hat{f}_{D_2}(t) = \left[\tilde{f}_{D_2}(t-1) \times 1.344170 - \tilde{f}_{D_2}(t-2) \times 1.332614 - \tilde{f}_{D_2}(t-3) \times 0.881539 - \tilde{f}_{D_2}(t-4) \times 0.324159 - \tilde{f}_{D_2}(t-28) \times 0.014090 \right] +$$

$$\begin{aligned} & [-\hat{e}_{D_2}(t-1) \times 0.907357 - \hat{e}_{D_2}(t-2) \times 0.447138 + \hat{e}_{D_2}(t-4) \times 0.149529 + \\ & \hat{e}_{D_2}(t-28) \times 0.006699] + \hat{\sigma}_{D_2,t}^2 \times 0. \end{aligned}$$

e o modelo EGARCH (1,1) e distribuição do erro generalizada (GED), é dado por:

$$\begin{aligned} \log(\hat{\sigma}_{D_3,t}^2) = & -9.084902 - \left| \left(\frac{\hat{e}_{D_3}(t-1)}{\sqrt{\hat{\sigma}_{D_3,t-1}^2}} \right) \right| \times 0.465595 + \left(\frac{\hat{e}_{D_3}(t-1)}{\sqrt{\hat{\sigma}_{D_3,t-1}^2}} \right) \times 0.042210 - \\ & \log(\hat{\sigma}_{D_3,t-1}^2) \times 0.178841. \end{aligned}$$

Dependent Variable: D3_F19X

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	-1.344170	0.021756	-61.78515	0.0000
AR(2)	-1.332614	0.026377	-50.52121	0.0000
AR(3)	-0.881539	0.028772	-30.63913	0.0000
AR(4)	-0.324159	0.017932	-18.07741	0.0000
AR(28)	0.014090	0.008610	1.636525	0.1017
MA(1)	-0.907357	0.033145	-27.37508	0.0000
MA(2)	-0.447138	6.56E-07	-681509.2	0.0000
MA(3)	0.207074	5.92E-05	3497.829	0.0000
MA(4)	0.149529	0.036533	4.093002	0.0000
MA(28)	0.006699	0.017152	0.390546	0.6961
Variance Equation				
C(11)	-9.084902	1.046544	-8.680863	0.0000
C(12)	0.465595	0.046912	9.924953	0.0000
C(13)	0.042210	0.025616	1.647782	0.0994
C(14)	0.178841	0.098655	1.812788	0.0699
GED PARAMETER	1.637199	0.078328	20.90179	0.0000
R-squared	0.925752	Mean dependent var		2.08E-06
Adjusted R-squared	0.925411	S.D. dependent var		0.021115
S.E. of regression	0.005767	Akaike info criterion		-7.636377
Sum squared resid	0.065015	Schwarz criterion		-7.593757
Log likelihood	7517.741	Hannan-Quinn criter.		-7.620714
Durbin-Watson stat	1.841971			

Heteroskedasticity Test: ARCH

F-statistic	0.000402	Prob. F(1,2482)	0.9840
Obs*R-squared	0.000402	Prob. Chi-Square(1)	0.9840

(3) Modelo ARIMA-GARCH para a componente de detalhe $\left(\tilde{f}_{D_3}(t)\right)_{t=1}^{1993}$, onde o modelo ARIMA é dado por:

$$\hat{f}_{D_3}(t) = [\tilde{f}_{D_3}(t-1) \times 0.874936 - \tilde{f}_{D_3}(t-2) \times 1.259681 + \tilde{f}_{D_3}(t-3) \times 0.513852 - \tilde{f}_{D_3}(t-4) \times 0.480487] + [\hat{e}_{D_3}(t-1) \times 0.414532 - \hat{e}_{D_3}(t-2) \times 0.670180 - \hat{e}_{D_3}(t-3) \times 0.227917 + \hat{e}_{D_3}(t-4) \times 0.150224]$$

e o modelo EGarch (1,1) e distribuição do erro generalizada (GED), é dado por:

$$\log(\hat{\sigma}_{D_3,t}^2) = -1.528652 - \left| \frac{\hat{e}_{D_3}(t-1)}{\sqrt{\hat{\sigma}_{D_3,t-1}^2}} \right| \times 0.492490 - \left(\frac{\hat{e}_{D_3}(t-1)}{\sqrt{\hat{\sigma}_{D_3,t-1}^2}} \right) \times 0.027170 - \log(\hat{\sigma}_{D_3,t-1}^2) \times 0.885977.$$

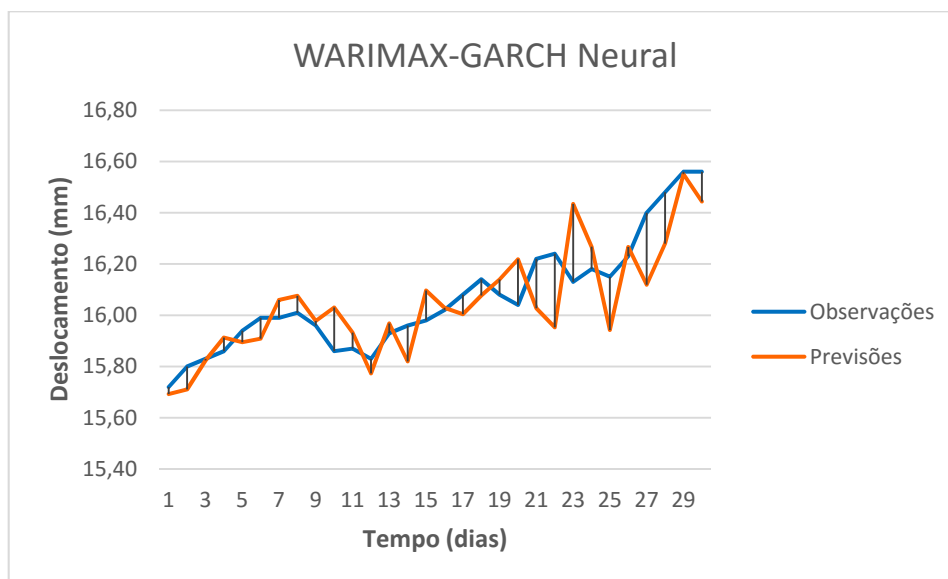
Dependent Variable: D3_F19X

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.874936	0.001089	803.1327	0.0000
AR(2)	-1.259681	0.005521	-228.1451	0.0000
AR(3)	0.513852	0.009586	53.60409	0.0000
AR(4)	-0.480487	0.008875	-54.13952	0.0000
MA(1)	0.414532	0.018674	22.19890	0.0000
MA(2)	-0.670180	0.005096	-131.5182	0.0000
MA(3)	-0.227917	0.025082	-9.086810	0.0000
MA(4)	-0.150224	0.011125	-13.50282	0.0000
Variance Equation				
C(9)	-1.528652	0.249053	-6.137851	0.0000
C(10)	0.492490	0.048766	10.09913	0.0000
C(11)	-0.027170	0.025326	-1.072842	0.2833
C(12)	0.885977	0.022382	39.58360	0.0000
GED PARAMETER	1.348534	0.065635	20.54611	0.0000
R-squared	0.956670	Mean dependent var	-1.51E-05	
Adjusted R-squared	0.956517	S.D. dependent var	0.033366	
S.E. of regression	0.006958	Akaike info criterion	-7.326698	
Sum squared resid	0.095901	Schwarz criterion	-7.290127	
Log likelihood	7299.401	Hannan-Quinn criter.	-7.313266	
Durbin-Watson stat	1.985216			

Apêndice V E: Evolução Temporal do Erro e Parâmetros do Modelo WARIMAX-GARCH Neural (F13y).

Na Figura 63, podemos comparar a evolução temporal das previsões fora da amostra para WARIMAX-GARCH Neural e ARIMA-GARCH em relação às observações de fora da amostra



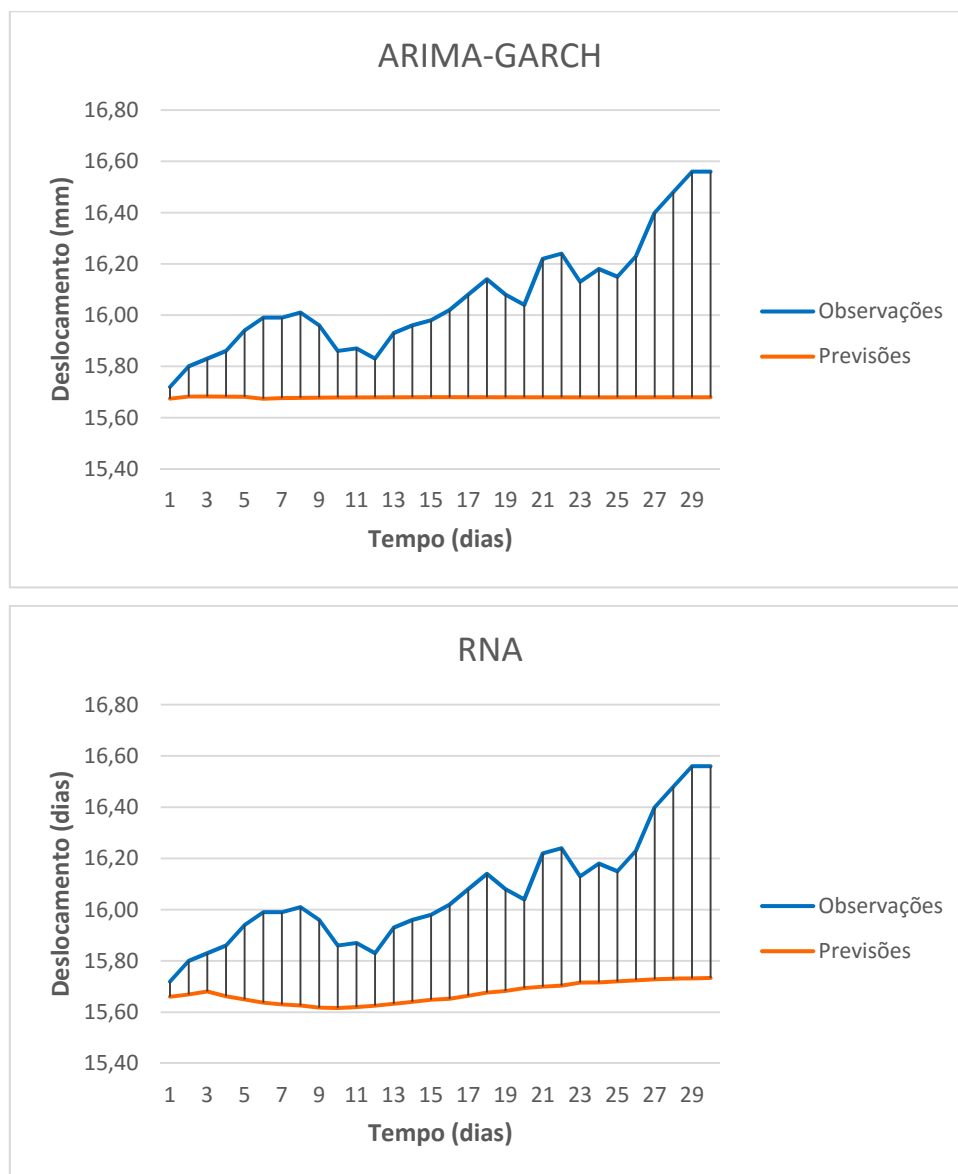


Figura 63 – Evolução temporal do erro APE para WARIMAX-GARCH Neural e ARIMA-GARCH e RNA

Dependent Variable: D(F19X)

Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
A2_F19XF(-1)	0.903728	0.012618	71.62361	0.0000
A2_F19XF(-2)	-0.903743	0.012616	-71.63468	0.0000
D1_F19XF(-1)	-1.999239	0.002800	-714.0183	0.0000
AR(1)	0.351502	0.024825	14.15910	0.0000
AR(7)	0.137590	0.022792	6.036668	0.0000
MA(2)	-0.865632	0.001904	-454.5525	0.0000
MA(1)	0.478306	0.009565	50.00438	0.0000
MA(3)	-0.376971	0.001576	-239.2170	0.0000

MA(19)	0.045470	0.010902	4.170813	0.0000
Variance Equation				
C	4.69E-05	1.74E-05	2.698024	0.0070
RESID(-1)^2	0.065001	0.017222	3.774216	0.0002
GARCH(-1)	0.873835	0.034621	25.23976	0.0000
GED PARAMETER	1.455074	0.061120	23.80693	0.0000
R-squared	0.842793	Mean dependent var	-0.000157	
Adjusted R-squared	0.841994	S.D. dependent var	0.069494	
S.E. of regression	0.027624	Akaike info criterion	-4.394367	
Sum squared resid	1.500233	Schwarz criterion	-4.351959	
Log likelihood	4358.832	Hannan-Quinn criter.	-4.378787	
Durbin-Watson stat	2.036502			

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 -0.008	-0.008	0.1142	
		2 -0.046	-0.046	4.3321	
		3 -0.022	-0.023	5.2901	
		4 -0.017	-0.020	5.8915	
		5 -0.008	-0.010	6.0103	
		6 0.016	0.013	6.4906	
		7 -0.001	-0.002	6.4915	
		8 -0.038	-0.038	9.3591	
		9 -0.013	-0.013	9.6864	0.011
		10 0.008	0.005	9.8182	0.013
		11 -0.016	-0.019	10.345	0.016
		12 0.017	0.015	10.904	0.028
		13 0.028	0.026	12.519	0.028
		14 0.001	0.003	12.520	0.051
		15 0.027	0.030	13.966	0.052
		16 -0.019	-0.019	14.708	0.065
		17 0.030	0.034	16.556	0.056
		18 0.018	0.019	17.216	0.070
		19 0.047	0.049	21.627	0.027
		20 -0.007	-0.002	21.716	0.041
		21 0.023	0.032	22.779	0.044
		22 0.012	0.017	23.079	0.059
		23 -0.015	-0.009	23.523	0.074
		24 0.013	0.016	23.847	0.093
		25 0.005	0.005	23.887	0.123
		26 -0.028	-0.023	25.466	0.113
		27 0.007	0.008	25.554	0.143
		28 0.072	0.072	36.013	0.015
		29 0.029	0.033	37.655	0.014
		30 0.002	0.009	37.661	0.020
		31 0.003	0.007	37.685	0.028
		32 0.045	0.048	41.756	0.014
		33 -0.020	-0.015	42.570	0.016
		34 0.033	0.030	44.803	0.012
		35 0.016	0.018	45.295	0.015
		36 -0.021	-0.016	46.226	0.017

FAC e FACP dos resíduos ordinários padronizados

Heteroskedasticity Test: ARCH

F-statistic	0.098281	Prob. F(1,1974)	0.7539
Obs*R-squared	0.098376	Prob. Chi-Square(1)	0.7538